

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2004, 00:06:13 ; Search time 145.506 Seconds
(without alignments)
11537.137 Million cell updates/sec

Title: US-10-626-173-1
Perfect score: 3025
Sequence: 1 gtcgaccacgctccggcc.....aaaaaaaaaaggcgccgc 3025

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCFUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	3025	100.0	3025	US-09-291-839-1
2	3025	100.0	3025	US-09-458-457-1
3	3025	100.0	3025	US-09-947-199A-1
4	2505	82.8	2505	US-09-291-839-3
5	2505	82.8	2505	US-09-458-457-3
6	2505	82.8	2505	US-09-947-199A-3
7	2010	66.4	3026	US-09-458-457-7
8	2010	66.4	3026	US-09-947-199A-7
9	1892.2	62.6	2505	US-09-458-457-9
10	1892.2	62.6	2505	US-09-947-199A-9
11	427.6	14.1	740	US-09-833-381-1151
12	427.4	14.1	616	US-09-833-381-1149
13	414.6	13.7	736	US-09-833-381-1150
14	184.2	6.1	304	US-09-833-381-1147
15	84.4	2.8	186	US-09-833-381-1148
16	84	2.8	14770	US-09-220-132-30
17	78.8	2.6	2894	US-09-620-312D-95
18	78.8	2.6	3606	US-10-164-595-37
19	78.6	2.6	3454	US-09-082-059-1
20	75.4	2.5	4297	US-09-972-115A-3
21	75.4	2.5	4493	US-09-972-115A-1
22	74.2	2.5	384	US-09-972-115A-9
23	74.2	2.5	1069	US-09-833-381-841
24	74.2	2.5	3394	US-09-427-154-1
25	74.2	2.5	3498	US-09-350-982C-4
26	74.2	2.5	3797	US-09-696-668-1
27	74.2	2.5	3816	US-09-696-668-2

28	74.2	2.5	4275	4	US-09-972-115A-5	Sequence 5, Appli
29	74.2	2.5	4512	4	US-09-350-982C-3	Sequence 3, Appli
30	72.8	2.4	1056	3	US-09-065-474-138	Sequence 138, App
31	72.8	2.4	1056	3	US-09-065-474-140	Sequence 140, App
32	72.8	2.4	1056	3	US-09-557-034-138	Sequence 138, App
33	72.8	2.4	1056	4	US-09-557-034-140	Sequence 140, App
34	72.8	2.4	5235	1	US-09-031-485-35	Sequence 35, Appl
35	72.8	2.4	5235	1	US-09-031-485-36	Sequence 36, Appl
36	72.8	2.4	5235	1	US-08-847-429A-35	Sequence 35, Appl
37	72.8	2.4	5235	1	US-08-847-429A-36	Sequence 36, Appl
38	72.8	2.4	5235	3	US-09-065-474-35	Sequence 35, Appl
39	72.8	2.4	5235	3	US-09-065-474-36	Sequence 36, Appl
40	72.8	2.4	5235	4	US-09-557-034-35	Sequence 35, Appl
41	72.8	2.4	5235	4	US-09-557-034-36	Sequence 36, Appl
42	72.8	2.4	5503	1	US-09-031-485-32	Sequence 32, Appl
43	72.8	2.4	5503	1	US-09-031-485-34	Sequence 34, Appl
44	72.8	2.4	5503	1	US-08-847-429A-32	Sequence 32, Appl
45	72.8	2.4	5503	1	US-08-847-429A-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-09-291-839-1
; Sequence 1, Application US/09291839A
; Patent No. 6261818
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-068
; CURRENT APPLICATION NUMBER: US/09/291.839A
; CURRENT FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2552)
US-09-291-839-1

Query Match		100.0%;	Score 3025;	DB 3;	Length 3025;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 3025;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GTGACCCACCGCTCCGCGCCCTGGAGAAAGGAGAAACTTATTAATCGGAAATTTATA	60		
DB	1	GTGACCCACCGCTCCGCGCCCTGGAGAAAGGAGAAACTTATTAATCGGAAATTTATA	60		
QY	61	AATCTAGACCAACCCAACTTGTACTCATGATGGAAGAAAGTCAAGTGAATCATATG	120		
DB	61	AATCTAGACCAACCCAACTTGTACTCATGATGGAAGAAAGTCAAGTGAATCATATG	120		
QY	121	TTATCACAATAGAAGATTAGAAGTCACTGCAGATCAAGGAAAGAACTGACAGAAC	180		
DB	121	TTATCACAATAGAAGATTAGAAGTCACTGCAGATCAAGGAAAGAACTGACAGAAC	180		
QY	181	TAAAGMATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTAAATTCGGCACTG	240		
DB	181	TAAAGMATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTAAATTCGGCACTG	240		
QY	241	AAATGGGCTGCTCTACTTTCATTTATGTTGCAATTTGTGAGGCAAGAAATCACATATTC	300		
DB	241	AAATGGGCTGCTCTACTTTCATTTATGTTGCAATTTGTGAGGCAAGAAATCACATATTC	300		
QY	301	GAATCTTTATTTGAAAGGGCTCGCCCATCTGCACTGACAGAATGTTTACAGCCT	360		
DB	301	GAATCTTTATTTGAAAGGGCTCGCCCATCTGCACTGACAGAATGTTTACAGCCT	360		
QY	361	TGCATTTAGCATTTTACAAAGGATAATGCAGAAATTGATCACTTCTCTCTCAGTGGAG	420		

Db 1201 GTTTGATGTGGCTTATGAAAAAGGCGATGATGCCATTGTCACACTCTCTGAAGCAATTATA 1260
QY 1261 AGAGACCAAGATGAATTCCTGCTTAATGAATATTCACAGCTGGAGGAGATGGCTCCT 1320
Db 1261 AGAGACCAAGATGAATTCCTGCTTAATGAATATTCACAGCTGGAGGAGATGGCTCCT 1320
QY 1321 ATGTGCTCTTCATCACCCTTTGGGGAAGATTAAGGATGACAAAAGAGAGAGGACAGATA 1380
Db 1321 ATGTGCTCTTCATCACCCTTTGGGGAAGATTAAGGATGACAAAAGAGAGAGGACAGATA 1380
QY 1381 TTCTCTCTTAAGAGCTGAATTCCTTCAATTCCTTCAGCTCTCAGAAATTCAGT 1440
Db 1381 TTCTCTCTTAAGAGCTGAATTCCTTCAATTCCTTCAGCTCTCAGAAATTCAGT 1440
QY 1441 TCCATCAGATTAATGCTCAGGTTCTTTTGGGAAGATATATAAGGACGATGACAGAAATA 1500
Db 1441 TCCATCAGATTAATGCTCAGGTTCTTTTGGGAAGATATATAAGGACGATGACAGAAATA 1500
QY 1501 AAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCAAGTCAGATGAGATA 1560
Db 1501 AAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCAAGTCAGATGAGATA 1560
QY 1561 TGTGTTGCCGAGAGGTGCCATTCTCTGCCAGCTCAATCATCCCTGGTAAATTCAGTTTG 1620
Db 1561 TGTGTTGCCGAGAGGTGCCATTCTCTGCCAGCTCAATCATCCCTGGTAAATTCAGTTTG 1620
QY 1621 TGGGTGCTTGGTGAATGATCCAGCCAGTTTGGCAATGTCACCTCAATACATATCAGGG 1680
Db 1621 TGGGTGCTTGGTGAATGATCCAGCCAGTTTGGCAATGTCACCTCAATACATATCAGGG 1680
QY 1681 GTTCTCTGTTCCCTCTTCATGAGCGAGAGAGATTCCTGATTTGGAGTCTAAATTA 1740
Db 1681 GTTCTCTGTTCCCTCTTCATGAGCGAGAGAGATTCCTGATTTGGAGTCTAAATTA 1740
QY 1741 TTATTGCACTAGATGTTGCCAAGGCGATGGAGTCTTCAAACTGACACAGCCAAATTA 1800
Db 1741 TTATTGCACTAGATGTTGCCAAGGCGATGGAGTCTTCAAACTGACACAGCCAAATTA 1800
QY 1801 TACATCGTGAATGAAAGTCAAAATATTCCTCTATGAGGATGGGATGCTGTGGTG 1860
Db 1801 TACATCGTGAATGAAAGTCAAAATATTCCTCTATGAGGATGGGATGCTGTGGTG 1860
QY 1861 CAGATTTTGGGAATCAGATTTCTACAGTCTCTGATGAAAGCAACATGACAAAACAC 1920
Db 1861 CAGATTTTGGGAATCAGATTTCTACAGTCTCTGATGAAAGCAACATGACAAAACAC 1920
QY 1921 CTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTACGCGAGTCACTCGGTACACCATCA 1980
Db 1921 CTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTACGCGAGTCACTCGGTACACCATCA 1980
QY 1981 AAGCAGATGCTTACGATGCTGTGCTGTGGGAATTCCTACCTGGGGAATTCAT 2040
Db 1981 AAGCAGATGCTTACGATGCTGTGCTGTGGGAATTCCTACCTGGGGAATTCAT 2040
QY 2041 TCGCTCATCTCAAGCAGCGGCTCGGCGAGCAGACATGGCTTACCACCAACATCAGACCTC 2100
Db 2041 TCGCTCATCTCAAGCAGCGGCTCGGCGAGCAGACATGGCTTACCACCAACATCAGACCTC 2100
QY 2101 CAAATTGGCTATTCATATCCAAAGCCCATATCATCTCTGATACGAGGGTGGAAACCAT 2160
Db 2101 CAAATTGGCTATTCATATCCAAAGCCCATATCATCTCTGATACGAGGGTGGAAACCAT 2160
QY 2161 GTCTTGAAGGAAGACCCGAATTTCTGAAGTTGTCATGAAGTTAGAGAGTGTCTGCA 2220
Db 2161 GTCTTGAAGGAAGACCCGAATTTCTGAAGTTGTCATGAAGTTAGAGAGTGTCTGCA 2220
QY 2221 ACATTGAGCTGATGCTCTCTGCATCAAGATAACAGAGTGGTCTCTCTCACCTTCTTCT 2280
Db 2221 ACATTGAGCTGATGCTCTCTGCATCAAGATAACAGAGTGGTCTCTCTCACCTTCTTCT 2280
QY 2281 CTTCTGATTTGCCCTGGTGAACCGGGGAGGACCTGGCCGAGTCAATGTCGAGCAATTAAGAA 2340
Db 2281 CTTCTGATTTGCCCTGGTGAACCGGGGAGGACCTGGCCGAGTCAATGTCGAGCAATTAAGAA 2340

QY 2341 GTCTGTTTCCGAATTCGAATATGCTCTAAATGCAAGGTCTCTATGCTGCTTGTCTCCAAAGTG 2400
Db 2341 GTCTGTTTCCGAATTCGAATATGCTCTAAATGCAAGGTCTCTATGCTGCTTGTCTCCAAAGTG 2400
QY 2401 CTGGACAATATTCCTCTCAAGGTCTGCTTTTGGAGGAGATGAAAAAGAGTCTTCAATACA 2460
Db 2401 CTGGACAATATTCCTCTCAAGGTCTGCTTTTGGAGGAGATGAAAAAGAGTCTTCAATACA 2460
QY 2461 CACCCATTGACAAATATGGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCATCTT 2520
Db 2461 CACCCATTGACAAATATGGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCATCTT 2520
QY 2521 GCCGAAATAGTAGCAGCTTTGAGGACAGCAGCTGACAGCATTCCGGCTATACCTAAGGAG 2580
Db 2521 GCCGAAATAGTAGCAGCTTTGAGGACAGCAGCTGACAGCATTCCGGCTATACCTAAGGAG 2580
QY 2581 AGTTTTCCTCCGAACTGACAGCAACGATTCACACCGCAAGCTGGCTTCCAATATA 2640
Db 2581 AGTTTTCCTCCGAACTGACAGCAACGATTCACACCGCAAGCTGGCTTCCAATATA 2640
QY 2641 ACATTTTACTCTCAAGGTCTCTTAAATTTGGGCTTGTTTTACTTGTCTTAAATTC 2700
Db 2641 ACATTTTACTCTCAAGGTCTCTTAAATTTGGGCTTGTTTTACTTGTCTTAAATTC 2700
QY 2701 CCCACTATTAGCAGCTTTGGATTTGCTTAAGGAATAATATGCAAAAAGAACCAAGACA 2760
Db 2701 CCCACTATTAGCAGCTTTGGATTTGCTTAAGGAATAATATGCAAAAAGAACCAAGACA 2760
QY 2761 GAATGTATGAAGATTTGTTTAAATTTGTAATTTAAATAAATAAATTTAGATCGTTACT 2820
Db 2761 GAATGTATGAAGATTTGTTTAAATTTGTAATTTAAATAAATAAATTTAGATCGTTACT 2820
QY 2821 TGGAAATGGAACCTTAAGTCTGTGTGACAGATAATAATATGCTTCTGGGCTGAAT 2880
Db 2821 TGGAAATGGAACCTTAAGTCTGTGTGACAGATAATAATATGCTTCTGGGCTGAAT 2880
QY 2881 ATGTAGATTTGTTTACAGCTATGGGTTTATTTCTAGAACATTTGTTCTTTTCTTTT 2940
Db 2881 ATGTAGATTTGTTTACAGCTATGGGTTTATTTCTAGAACATTTGTTCTTTTCTTTT 2940
QY 2941 CTCATTATGTTTACTTCTAGTGTTCACCTCTGTGATTAAGATTTCTTTGGTGAATAGAAA 3000
Db 2941 CTCATTATGTTTACTTCTAGTGTTCACCTCTGTGATTAAGATTTCTTTGGTGAATAGAAA 3000
QY 3001 AAAAAAAGGAGGCGGCGCGC 3025
Db 3001 AAAAAAAGGAGGCGGCGCGC 3025

RESULT 3

US-09-947-199A-1

; Sequence 1, Application US/09947199A

; Patent No. 6660490

; GENERAL INFORMATION:

; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES

; FILE REFERENCE: MN1-068CP2

; CURRENT APPLICATION NUMBER: US/09/947,199A

; PRIOR FILING DATE: 2001-09-05

; PRIOR APPLICATION NUMBER: US 60/111,938

; PRIOR FILING DATE: 1998-12-11

; PRIOR APPLICATION NUMBER: US 09/291,839

; PRIOR FILING DATE: 1999-04-14

; PRIOR APPLICATION NUMBER: US 09/458,457

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 1

; LENGTH: 3025

; TYPE: DNA

; ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2552)
US-09-947-199A-1

Query Match 100.0%; Score 3025; DB 4; Length 3025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTCGACCCAGCGTCCGGCCCTGGAGAAAGGAGAAACTTATAATTAATGGGAATTTATA 60
Db 1 GTCGACCCAGCGTCCGGCCCTGGAGAAAGGAGAAACTTATAATTAATGGGAATTTATA 60
Qy 61 AATCTAGACCAACCCAAACTTGTACTGATGAATGGAAGAAAAGTCAGTGAATCATATG 120
Db 61 AATCTAGACCAACCCAAACTTGTACTGATGAATGGAAGAAAAGTCAGTGAATCATATG 120
Qy 121 TTATCACAATAGAAAGATTAAGATGACCTGAGATCAAGGAAAAGAACTGACAGAAC 180
Db 121 TTATCACAATAGAAAGATTAAGATGACCTGAGATCAAGGAAAAGAACTGACAGAAC 180
Qy 181 TAAGGAATATATTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTAAATTAACCGCAGT 240
Db 181 TAAGGAATATATTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTAAATTAACCGCAGT 240
Qy 241 AAAATGGGCTGTCTACTTCAATTTATGTTGCAATTTGTGAGGCAAGAAATCACATATTC 300
Db 241 AAAATGGGCTGTCTACTTCAATTTATGTTGCAATTTGTGAGGCAAGAAATCACATATTC 300
Qy 301 GAACCTTATGTTGAAGGGCTCGGCCATCTCGACTGACAGAAATGATTAAGCCT 360
Db 301 GAACCTTATGTTGAAGGGCTCGGCCATCTCGACTGACAGAAATGATTAAGCCT 360
Qy 361 TGCAATTTAGCAGTTTACAAGGAAATGCAAGAAATGATCACTTCTGCTTTCACAGTGGAG 420
Db 361 TGCAATTTAGCAGTTTACAAGGAAATGCAAGAAATGATCACTTCTGCTTTCACAGTGGAG 420
Qy 421 CTGATATACAGAGGTTGGATACGGTGGCCCTCACTGCCCTCCATATGCTACAAATGCTG 480
Db 421 CTGATATACAGAGGTTGGATACGGTGGCCCTCACTGCCCTCCATATGCTACAAATGCTG 480
Qy 481 GCCACCTTAGAGGCTGCTGATGCTGTTGCAACATGGAGCTAATGCAATATTAAGATG 540
Db 481 GCCACCTTAGAGGCTGCTGATGCTGTTGCAACATGGAGCTAATGCTAATATTAAGATG 540
Qy 541 CAGTTTTTTTCACTCCATTCGATATTCAGCGTACTATGGACATGAACAGGTAACCTCGCC 600
Db 541 CAGTTTTTTTCACTCCATTCGATATTCAGCGTACTATGGACATGAACAGGTAACCTCGCC 600
Qy 601 TTCTTTTGAATTTGGTCTGATGCTAAATGCTAAGTGGTGAAGTTGAGATGACCCCTCC 660
Db 601 TTCTTTTGAATTTGGTCTGATGCTAAATGCTAAGTGGTGAAGTTGAGATGACCCCTCC 660
Qy 661 ACCTAGCATCTGCAAAAAGGATTTCTTGAATATTCGAAAATCTTTGATGGAAGAGCAGCA 720
Db 661 ACCTAGCATCTGCAAAAAGGATTTCTTGAATATTCGAAAATCTTTGATGGAAGAGCAGCA 720
Qy 721 AAGCAGATGTAATGCTCAAGATTAAGAGACCATGTCCTCACTTCTGTTCTCGAT 780
Db 721 AAGCAGATGTAATGCTCAAGATTAAGAGACCATGTCCTCACTTCTGTTCTCGAT 780
Qy 781 TTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACCTCATG 840
Db 781 TTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACCTCATG 840
Qy 841 TTGTTAATATCTATGGAAGATACCCCTTACCTGGGATGCTCAATGCAAAATTTGAAG 900
Db 841 TTGTTAATATCTATGGAAGATACCCCTTACCTGGGATGCTCAATGCAAAATTTGAAG 900
Qy 901 TTGCGAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTRAGGAAACATCTTCA 960
Db 901 TTGCGAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTRAGGAAACATCTTCA 960

Qy 961 GTGAAACAGCTTTTCATAGTGTCTTACCTATGCGAAGAGCATTGACCTAGTCAAAATTC 1020
Db 961 GTGAAACAGCTTTTCATAGTGTCTTACCTATGCGAAGAGCATTGACCTAGTCAAAATTC 1020
Qy 1021 TTCTTGATCAGAAATGCTATAAACAATCAACCAAGGAAGGATGGGCACACTGGATTAC 1080
Db 1021 TTCTTGATCAGAAATGCTATAAACAATCAACCAAGGAAGGATGGGCACACTGGATTAC 1080
Qy 1081 ACTCTGCTTCTACACAGGTCACATTCGGCTCGTTCAGTTCCTTACTGGGATATGAGCTG 1140
Db 1081 ACTCTGCTTCTACACAGGTCACATTCGGCTCGTTCAGTTCCTTACTGGGATATGAGCTG 1140
Qy 1141 ATATGAATCTAGTGGCTTGTGATCCACAGAGTCTAGTGGTGAAGAAAGATGAGCAGAT 1200
Db 1141 ATATGAATCTAGTGGCTTGTGATCCACAGAGTCTAGTGGTGAAGAAAGATGAGCAGAT 1200
Qy 1201 GTTGTGATGGGCTTATGAAGAAGGATGATGCCATTCGTACACTCCTGAAGCATATA 1260
Db 1201 GTTGTGATGGGCTTATGAAGAAGGATGATGCCATTCGTACACTCCTGAAGCATATA 1260
Qy 1261 AGAGACCAACAAGATGAATTCGCCCTGTAATGAATATTTCTCAGCCTGGAGGATGGCTCT 1320
Db 1261 AGAGACCAACAAGATGAATTCGCCCTGTAATGAATATTTCTCAGCCTGGAGGATGGCTCT 1320
Qy 1321 ATGTGCTGTTCATCACTCCCTTGGGGAAGATTAAGAAGCATGACAAAAGAGAGGAGATA 1380
Db 1321 ATGTGCTGTTCATCACTCCCTTGGGGAAGATTAAGAAGCATGACAAAAGAGAGGAGATA 1380
Qy 1381 TTCTCCTCCTAAAGAGTGGATTCGCTTCACTTTCCATCTTCAGCTCTCAGAAATTTGAGT 1440
Db 1381 TTCTCCTCCTAAAGAGTGGATTCGCTTCACTTTCCATCTTCAGCTCTCAGAAATTTGAGT 1440
Qy 1441 TCCATGAGATTAATGGCTCAGGTTCTTTTGGGAAAGTATATAAAGGACGATCAGAAATA 1500
Db 1441 TCCATGAGATTAATGGCTCAGGTTCTTTTGGGAAAGTATATAAAGGACGATCAGAAATA 1500
Qy 1501 AAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCCAGTCAGATGGGATA 1560
Db 1501 AAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCCAGTCAGATGGGATA 1560
Qy 1561 TGTTTTGCCGAGAGGTGTCCATTTCTGCGCAGCTCAATCATCCCTGCGTAATTCAGTTTG 1620
Db 1561 TGTTTTGCCGAGAGGTGTCCATTTCTGCGCAGCTCAATCATCCCTGCGTAATTCAGTTTG 1620
Qy 1621 TGGGTGCTTCTGTAATGATCCGAGCAGTTTGCCATTTGTCATCAATACATATCAGGG 1680
Db 1621 TGGGTGCTTCTGTAATGATCCGAGCAGTTTGCCATTTGTCATCAATACATATCAGGG 1680
Qy 1681 GTTCTCTGTTCTCCCTCTCATGACAGAGAGGATTTCTGATTTGCACTCTAAATTA 1740
Db 1681 GTTCTCTGTTCTCCCTCTCATGACAGAGAGGATTTCTGATTTGCACTCTAAATTA 1740
Qy 1741 TTATTGCAATGATGTCGCAAGGATGAGGATGAGTACCTTCAACCTGACACAGCCAAATTA 1800
Db 1741 TTATTGCAATGATGTCGCAAGGATGAGGATGAGTACCTTCAACCTGACACAGCCAAATTA 1800
Qy 1801 TACATCGTGAATGAAACAGTCAATATTTCTCTATGAGGATGGGATGCTGTTGGTGG 1860
Db 1801 TACATCGTGAATGAAACAGTCAATATTTCTCTATGAGGATGGGATGCTGTTGGTGG 1860
Qy 1861 CAGATTTTGAGGAATCAAGATTTTCTACGTCTCTGATGAAGACAAACATGACAAAACAAC 1920
Db 1861 CAGATTTTGAGGAATCAAGATTTTCTACGTCTCTGATGAAGACAAACATGACAAAACAAC 1920
Qy 1921 CTGGGAACCTCCGTTGGATGGCTCTGAGGTTGTTACGCGAGTGCATCGGTACACATCA 1980
Db 1921 CTGGGAACCTCCGTTGGATGGCTCTGAGGTTGTTACGCGAGTGCATCGGTACACATCA 1980
Qy 1981 AAGCAGATGCTTTCAGCTATGCTGTGTGTTGGGAAATTTCTCACTGGCGAAATTTCCAT 2040
Db 1981 AAGCAGATGCTTTCAGCTATGCTGTGTGTTGGGAAATTTCTCACTGGCGAAATTTCCAT 2040
Qy 2041 TCGCTCATCTCAAGCCAGCGGCTGCGGCGACGAGATGCGGTTTACCAACATCAGACCTC 2100

```
Db 2041 TCCTCATCTCAAGCCAGGCTGGGAGCAGACATGGCTTACCACACATCAGACCTC 2100
QY 2101 CCATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATACGAGGGTGAACGCAT 2160
Db 2101 CCATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATACGAGGGTGAACGCAT 2160
QY 2161 GTCTCTGAAGAAAGACCCGAAATTTCTGAAGTGTTCATGAAGTTAGAAGAGTGTCTTGCA 2220
Db 2161 GTCTCTGAAGAAAGACCCGAAATTTCTGAAGTGTTCATGAAGTTAGAAGAGTGTCTTGCA 2220
QY 2221 ACATTGAGCTGATGTCCTCTGCAATCAAGTAACAGAGTGGTCTCTCTCACTTCTTCTT 2280
Db 2221 ACATTGAGCTGATGTCCTCTGCAATCAAGTAACAGAGTGGTCTCTCTCACTTCTTCTT 2280
QY 2281 CTCTGATTTGCTGTGAACCGGGAGGAGCTGCGCGAGTCAATGGCAGCATTAAGAA 2340
Db 2281 CTCTGATTTGCTGTGAACCGGGAGGAGCTGCGCGAGTCAATGGCAGCATTAAGAA 2340
QY 2341 GTCTGTTGCAATTTGAATATGCTCTAAATGCAAGTCTATGCTGCTTTGTGCCAAGTG 2400
Db 2341 GTCTGTTGCAATTTGAATATGCTCTAAATGCAAGTCTATGCTGCTTTGTGCCAAGTG 2400
QY 2401 CTGGCAATATTCTCTCAAGTCTGCTTTGGAGGAGATGAAAGAGTCTTCAATACA 2460
Db 2401 CTGGCAATATTCTCTCAAGTCTGCTTTGGAGGAGATGAAAGAGTCTTCAATACA 2460
QY 2461 CACCCATTGACAAATATGGCTATGTATCCGATCCCATGAGCTCAATGCAATTTTCATCTT 2520
Db 2461 CACCCATTGACAAATATGGCTATGTATCCGATCCCATGAGCTCAATGCAATTTTCATCTT 2520
QY 2521 GCGGAAATAGTAGAGCTTTGAGGACAGCAGCTGACAGCATTCGGCGTATACCTTAGGAG 2580
Db 2521 GCGGAAATAGTAGAGCTTTGAGGACAGCAGCTGACAGCATTCGGCGTATACCTTAGGAG 2580
QY 2581 AGTTTTTTCCCGAACTGACAGCAACGATTCACACCGGCAAGCTGGCTTCCAACTATA 2640
Db 2581 AGTTTTTTCCCGAACTGACAGCAACGATTCACACCGGCAAGCTGGCTTCCAACTATA 2640
QY 2641 ACATTTACTCTCAAGGTCTCTTAAATGGGCTTGTTTTACTGCTCTTAAATTC 2700
Db 2641 ACATTTACTCTCAAGGTCTCTTAAATGGGCTTGTTTTACTGCTCTTAAATTC 2700
QY 2701 CCCACTATTAGCAGCTTTGGATTTGTCCTAAGGAAATATATGCAAAAGAACCAAGACA 2760
Db 2701 CCCACTATTAGCAGCTTTGGATTTGTCCTAAGGAAATATATGCAAAAGAACCAAGACA 2760
QY 2761 GAATGTATATGAAGAAATTTGTTTTAAATTTCTAAATTTAAATAAATAAATAAATAAATA 2820
Db 2761 GAATGTATATGAAGAAATTTGTTTTAAATTTCTAAATTTAAATAAATAAATAAATAAATA 2820
QY 2821 TGGAAATGGAGCTTAAGTCTGTGGGAGCAGATTAATTAATTTCTTGGGCTGAAT 2880
Db 2821 TGGAAATGGAGCTTAAGTCTGTGGGAGCAGATTAATTAATTTCTTGGGCTGAAT 2880
QY 2881 ATGTAGACTGTGTTTGAAGCTATGGGTTTATTTCTTAGAAATTTCTTCTTCTTCTT 2940
Db 2881 ATGTAGACTGTGTTTGAAGCTATGGGTTTATTTCTTAGAAATTTCTTCTTCTTCTTCTT 2940
QY 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGATTAAGATTTCTTGGTGAATAGAAA 3000
Db 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGATTAAGATTTCTTGGTGAATAGAAA 3000
QY 3001 AAAAAAAAAAAAAAAAAAGGCGCGC 3025
Db 3001 AAAAAAAAAAAAAAAAAAGGCGCGC 3025
```

RESULT 4
US-09-291-839-3
; Sequence 3, Application US/09291839A
; Patent No. 6261818
; GENERAL INFORMATION:

```
; APPLICANT: Raju, Jeyaseelan  
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREOF  
; FILE REFERENCE: MNI-068  
; CURRENT APPLICATION NUMBER: US/09/291,839A  
; CURRENT FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2505  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2505)  
US-09-291-839-3
```

Query Match 82.8%; Score 2505; DB 3; Length 2505;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 48 ATGGAAATTAATAATCTAGACCAACCCAAACTTGTACTGATGAATGGAGAAAAGTTC 107  
Db 1 ATGGAAATTAATAATCTAGACCAACCCAAACTTGTACTGATGAATGGAGAAAAGTTC 60  
QY 108 AGTGAATCATATGTTATCAANTAGAAAGATTAGAAGTACCTGCAGATCAAGGAAAA 167  
Db 61 AGTGAATCATATGTTATCAANTAGAAAGATTAGAAGTACCTGCAGATCAAGGAAAA 120  
QY 168 GAACTGACAGAACTTAAGGAATATATTGGCTCTGTAGTGAAGCCTTCAGTAAAGTCAATT 227  
Db 121 GAACTGACAGAACTTAAGGAATATATTGGCTCTGTAGTGAAGCCTTCAGTAAAGTCAATT 180  
QY 228 AATTACGCGACTGAAATGGGCTGCTCTACTTCACTTATGTTGCAATTTGTGGAGGCAAG 287  
Db 181 AATTACGCGACTGAAATGGGCTGCTCTACTTCACTTATGTTGCAATTTGTGGAGGCAAG 240  
QY 288 AAATCAATATTCGAACCTTTATGTTGAAAGGCTCCGCCCATCTCGACTGACAAAGAAAT 347  
Db 241 AAATCAATATTCGAACCTTTATGTTGAAAGGCTCCGCCCATCTCGACTGACAAAGAAAT 300  
QY 348 GGATTTACAGCTTTGCAATTTAGCAGTTTACAGGAATAATGCAAGATTGATCTCTCTG 407  
Db 301 GGATTTACAGCTTTGCAATTTAGCAGTTTACAGGAATAATGCAAGATTGATCTCTCTG 360  
QY 408 CTTCACAGTGGAGCTGATATACAGAGTGTGGATACGGTGGCTCACTCCCTCCATATT 467  
Db 361 CTTCACAGTGGAGCTGATATACAGAGTGTGGATACGGTGGCTCACTCCCTCCATATT 420  
QY 468 GCTACAATAGCTGGCCACCTAGAGGCTGTGATGCTGTGCAACATGGAGCTAATGTC 527  
Db 421 GCTACAATAGCTGGCCACCTAGAGGCTGTGATGCTGTGCAACATGGAGCTAATGTC 480  
QY 528 AATATTCAAGATCAGATTTTTTTTCACTCCATTCGATATTCGACGCTACTATGACATGAA 587  
Db 481 AATATTCAAGATCAGATTTTTTTTCACTCCATTCGATATTCGACGCTACTATGACATGAA 540  
QY 588 CAGTAACTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTCGA 647  
Db 541 CAGTAACTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTCGA 600  
QY 648 GATAGACCCCTCCACCTAGCATCTGAAAGAGTCTTGAATTTGCAAACTCTTGATG 707  
Db 601 GATAGACCCCTCCACCTAGCATCTGAAAGAGTCTTGAATTTGCAAACTCTTGATG 660  
QY 708 GAAGAAGCAGCAAGCAGATGTGAATGTCTCAAGATAATCAAGACCATCTCCACTCCAT 767  
Db 661 GAAGAAGCAGCAAGCAGATGTGAATGTCTCAAGATAATCAAGACCATCTCCACTCCAT 720  
QY 768 TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATGGAA 827  
Db 721 TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATGGAA 780  
QY 828 GTTCAACCTCATGTGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAA 887
```

Db 781 |||||GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACACCTGCGATGCTACAAAT 840
Qy 888 GGCNAATTTGAAGTTGCCAGGAATATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947
Db 841 GGCNAATTTGAAGTTGCCAGGAATATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
Qy 948 GAAAAATCTTCAAGTGAACAGCTTTTTCATAGTGTGTTACCTATGCGCAAGAGCAATTCAC 1007
Db 901 GAAAAATCTTCAAGTGAACAGCTTTTTCATAGTGTGTTACCTATGCGCAAGAGCAATTCAC 960
Qy 1008 CTAGTCAAAATTTCTTGTGATCAGATGTCATTAACATCAACACCAAGGAAGGATGGG 1067
Db 961 CTAGTCAAAATTTCTTGTGATCAGATGTCATTAACATCAACACCAAGGAAGGATGGG 1020
Qy 1068 CACACTGGATTTACACTCTGCTTCTACCAAGGTCACATTCGCTGTGTTTCAAGTTCCTTACTG 1127
Db 1021 CACACTGGATTTACACTCTGCTTCTACCAAGGTCACATTCGCTGTGTTTCAAGTTCCTTACTG 1080
Qy 1128 GATAATGGAGCTGATATGAATATAGTGGCTTGTGATCCAGCAGGTCATAGTGGTGAAGAAA 1187
Db 1081 GATAATGGAGCTGATATGAATATAGTGGCTTGTGATCCAGCAGGTCATAGTGGTGAAGAAA 1140
Qy 1188 GATGAGCAGACATGTTTGTATGGGCTTATGAAGAGGCGATGATGCCATTTGTACACTC 1247
Db 1141 GATGAGCAGACATGTTTGTATGGGCTTATGAAGAGGCGATGATGCCATTTGTACACTC 1200
Qy 1248 CTGAAGCAATTTAAGAGACACCAAGATGAATTCGCTGTGAATTAATTTCTCAGGCTCGA 1307
Db 1201 CTGAAGCAATTTAAGAGACACCAAGATGAATTCGCTGTGAATTAATTTCTCAGGCTCGA 1260
Qy 1308 GGAGATGGCTCCTATGATGCTGTGTTTCCATCACCCTTGGGGAAGATTAAGAGCATGCAAAA 1367
Db 1261 GGAGATGGCTCCTATGATGCTGTGTTTCCATCACCCTTGGGGAAGATTAAGAGCATGCAAAA 1320
Qy 1368 GAGAAGCAGATATTCCTCCTTAAGAGCTGGATTCGCTTCAATTTCCATCTTCAGCTC 1427
Db 1321 GAGNAGCAGATATTCCTCCTTAAGAGCTGGATTCGCTTCAATTTCCATCTTCAGCTC 1380
Qy 1428 TCAGAAATTCAGTTCATGAGATATTTGCTCAGGTTCTTTTGGGAAGATATATAAGGA 1487
Db 1381 TCAGAAATTCAGTTCATGAGATATTTGCTCAGGTTCTTTTGGGAAGATATATAAGGA 1440
Qy 1488 CGATGCAGAAATAAATAGTGGCTATTAACAGTTATCGAGCCAAATACCTACTGCTCCCAAG 1547
Db 1441 CGATGCAGAAATAAATAGTGGCTATTAACAGTTATCGAGCCAAATACCTACTGCTCCCAAG 1500
Qy 1548 TCAGATGTGATATGTTTTGCCGAGAGGTGTCCATTTCTTGCAGCTCAATCATCCCTGC 1607
Db 1501 TCAGATGTGATATGTTTTGCCGAGAGGTGTCCATTTCTTGCAGCTCAATCATCCCTGC 1560
Qy 1608 GTAAATTCAGTTTGGGTGCTTCTGTTGAATGATCCAGCAGTTTGCCTATGCTACTCAA 1667
Db 1561 GTAAATTCAGTTTGGGTGCTTCTGTTGAATGATCCAGCAGTTTGCCTATGCTACTCAA 1620
Qy 1668 TACATATCAGGGGTTCTCTGTTCTCCCTCATGAGCAGAAAGAGGATTTCTTGATTTG 1727
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCATGAGCAGAAAGAGGATTTCTTGATTTG 1680
Qy 1728 CAGTCTAAATTAATATTGAGTATGTTGCAAGGATGAGTACCTTCACAACTG 1787
Db 1681 CAGTCTAAATTAATATTGAGTATGTTGCAAGGATGAGTACCTTCACAACTG 1740
Qy 1788 ACACAGCAATATATACATCTGACTGTAAGCAGTCACAAATATTTCTCTATGAGATGGG 1847
Db 1741 ACACAGCAATATATACATCTGACTGTAAGCAGTCACAAATATTTCTCTATGAGATGGG 1800
Qy 1848 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1907
Db 1801 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1860
Qy 1908 ATGCAAAAAACCTGGAACTCCGTTGGATGGCTTCTGAGGTGTTCCAGCTGCACT 1967

Db 1861 ATGACAAAAAACCTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTTCAACGAGTGCAC 1920
Qy 1968 CGGTACACCATCAAGCAGATGTTTCAAGTATGCTCTGTGCTGTGGAAATTTCTCACT 2027
Db 1921 CGGTACACCATCAAGCAGATGTTTCAAGTATGCTCTGTGCTGTGGAAATTTCTCACT 1980
Qy 2028 GCGAAATTTCCATTCGCTCATCTCAAGCAGCGGTGCGGACAGCAGATGGCTTACCAC 2087
Db 1981 GCGAAATTTCCATTCGCTCATCTCAAGCAGCGGTGCGGACAGCAGATGGCTTACCAC 2040
Qy 2088 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2147
Db 2041 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2100
Qy 2148 GGGTGAACGCAATGCTCTGAAGGAAGACCCGAATTTCTGAAGTTGTCTATGAAGTTAGAA 2207
Db 2101 GGGTGAACGCAATGCTCTGAAGGAAGACCCGAATTTCTGAAGTTGTCTATGAAGTTAGAA 2160
Qy 2208 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTGTCATCAAGTAAACAGAGTGGGTCTCTC 2267
Db 2161 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTGTCATCAAGTAAACAGAGTGGGTCTCTC 2220
Qy 2268 TCACCTTCTTCTTCTTCTGATTGCTGTAACCGGGGAGGACCTGGCGGAGTCACTGTG 2327
Db 2221 TCACCTTCTTCTTCTGATTGCTGTAACCGGGGAGGACCTGGCGGAGTCACTGTG 2280
Qy 2328 GAGCATTTAAGAAAGTCTGTTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2387
Db 2281 GCAGCATTTAAGAAAGTCTGTTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2340
Qy 2388 TTGTCCCAAGTCTCGCAATATTTCTCTCAAGTCTGCTTCTTGAGAGATGAAAGA 2447
Db 2341 TTGTCCCAAGTCTCGCAATATTTCTCTCAAGTCTGCTTCTTGAGAGATGAAAGA 2400
Qy 2448 AGTCTTCAATACACACCCATTCGAAATATGCTATGATATCCGATCCCATGAGCTCAATG 2507
Db 2401 AGTCTTCAATACACACCCATTCGAAATATGCTATGATATCCGATCCCATGAGCTCAATG 2460
Qy 2508 CATTTTCATTTGCGGAATATGATAGCAGCTTTTGAGGACAGCAGC 2552
Db 2461 CATTTTCATTTGCGGAATATGATAGCAGCTTTTGAGGACAGCAGC 2505

RESULT 5

US-09-458-457-3
; Sequence 3, Application US/09458457
; Patent No. 650654
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: nri-068CD
; CURRENT APPLICATION NUMBER: US/09/458,457
; EARLIER FILING DATE: 1999-12-10
; EARLIER APPLICATION NUMBER: 60/111,938
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/291,839
; EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-458-457-3

Query Match 82.8%; Score 2505; DB 4; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	48	ATGGGAATTAATAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAGTC	107		1081	GATTAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCAGCAGGCTCTAGTGGTGA	1140
DB	1	ATGGGAATTAATAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAGTC	60		1188	GATCAGCAGACATGTTTGTATGGGCTTATGAAAAAGGGCATGATGCCATTGTACACTC	1247
QY	108	AGTGAATCATATGTTTATCACAAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA	167		1141	GATCAGCAGACATGTTTGTATGGGCTTATGAAAAAGGGCATGATGCCATTGTACACTC	1200
DB	61	AGTGAATCATATGTTTATCACAAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA	120		1248	CTGAAGCATTTATTAAGAGACCAACAAGATGAATTCGCCCTGTAAATGAAATTTCTCAGCTCGA	1307
QY	168	GAACTGACAGAACTAAGGAATATATTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA	227		1201	CTGAAGCATTTATTAAGAGACCAACAAGATGAATTCGCCCTGTAAATGAAATTTCTCAGCTCGA	1260
DB	121	GAACTGACAGAACTAAGGAATATATTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA	180		1308	GGAGATGGCTCTTATGTTGTTCCATCACCCCTTGGGAAGATTAAAAAGCATGACAAA	1367
QY	228	AATTACCGCACTGAAATAGGGCTGCTCTACTTCATTTATGTTGCAATTTGTGAGGCAAG	287		1261	GGAGATGGCTCTTATGTTGTTCCATCACCCCTTGGGAAGATTAAAAAGCATGACAAA	1320
DB	181	AATTACCGCACTGAAATAGGGCTGCTCTACTTCATTTATGTTGCAATTTGTGAGGCAAG	240		1368	GAGAAGCAGATATTCCTCCTAAAGCTGGAATTCCTTCAATTTCCATTTCCATCTCAGCTC	1427
QY	288	AAATCACATATTCGAACTCTTATGTTGAAAGGGCTCCGCCCATCTCGACTGACAAAGAAAT	347		1321	GAGAAGCAGATATTCCTCCTAAAGCTGGAATTCCTTCAATTTCCATTTCCATCTCAGCTC	1380
DB	241	AAATCACATATTCGAACTCTTATGTTGAAAGGGCTCCGCCCATCTCGACTGACAAAGAAAT	300		1428	TCAGAAATTCAGATTCATGAGATTAATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA	1487
QY	348	GGATTTACAGCCTTCGATTTAGCAGTTTACAGGATTAATCAGAAATTCATCTCTCTG	407		1381	TCAGAAATTCAGATTCATGAGATTAATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA	1440
DB	301	GGATTTACAGCCTTCGATTTAGCAGTTTACAGGATTAATCAGAAATTCATCTCTCTG	360		1488	CGATGCAAAATAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTTACTGCTCCAG	1547
QY	408	CTTCAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTCACTGCCCTCCATATT	467		1441	CGATGCAAAATAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTTACTGCTCCAG	1500
DB	361	CTTCAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTCACTGCCCTCCATATT	420		1548	TCAGATGCGATATGTTTTGCCGAGAGGTGTCATCTCTGCGAGCTCAATCATCCCTGC	1607
QY	468	GCTAATAAGTGGCCACTAGAGGCTGCTGATGCTGTTGCAACATGAGCTAATGTC	527		1501	TCAGATGCGATATGTTTTGCCGAGAGGTGTCATCTCTGCGAGCTCAATCATCCCTGC	1560
DB	421	GCTAATAAGTGGCCACTAGAGGCTGCTGATGCTGTTGCAACATGAGCTAATGTC	480		1608	GTAATTCAGTTGTTGGTGTCTTGAATGATCCAGCAGTTTGGCCATTGTCTACTCAA	1667
QY	528	AATATTCAAGATGCAATTTTTTTCACTCCATTCGATATTCAGCGTACTATGACATGAA	587		1561	GTAATTCAGTTGTTGGTGTCTTGAATGATCCAGCAGTTTGGCCATTGTCTACTCAA	1620
DB	481	AATATTCAAGATGCAATTTTTTTCACTCCATTCGATATTCAGCGTACTATGACATGAA	540		1668	TACATATCAGGGGTTCTCTGTTCTCCCTTCATGAGCAGAAGAGGATTCCTGATTG	1727
QY	588	CAGGTAATCTGGCTTTTGAATTTGGTCTGATGTAATGTAAGTGTGAAGTTGGA	647		1621	TACATATCAGGGGTTCTCTGTTCTCCCTTCATGAGCAGAAGAGGATTCCTGATTG	1680
DB	541	CAGGTAATCTGGCTTTTGAATTTGGTCTGATGTAATGTAAGTGTGAAGTTGGA	600		1728	CAGTCTAAATTAATTTGCAAGTAGATTTGCCAAAGGCATGGAGTACCTTCAACACCTG	1787
QY	648	GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATGCAAACTCTTGATG	707		1681	CAGTCTAAATTAATTTGCAAGTAGATTTGCCAAAGGCATGGAGTACCTTCAACACCTG	1740
DB	601	GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATGCAAACTCTTGATG	660		1788	ACAGCCCAATTAATGATCGTGACTTGAACAGTCACAATATTTCTCTCTATCAGGATGG	1847
QY	708	GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGATTAATGAAGCAGATGCCACCTCCAT	767		1741	ACAGCCCAATTAATGATCGTGACTTGAACAGTCACAATATTTCTCTCTATCAGGATGG	1800
DB	661	GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGATTAATGAAGCAGATGCCACCTCCAT	720		1848	CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGATGAAGACACAC	1907
QY	768	TTCTGTTCTCGAATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTTGGA	827		1801	CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGATGAAGACACAC	1860
DB	721	TTCTGTTCTCGAATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTTGGA	780		1908	ATGACAAAAACACTGGAACTCCGTTGGATGGCTCTCAGGTTTACGAGTGTTCAGCAGTGCAT	1967
QY	828	TTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTTACACTGGCATGCTCAAT	887		1861	ATGACAAAAACACTGGAACTCCGTTGGATGGCTCTCAGGTTTACGAGTGTTCAGCAGTGCAT	1920
DB	781	TTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTTACACTGGCATGCTCAAT	840		1968	CGGTACACCATCAAGCAGATGTTTCAAGTATGCTCTGTCTGTGGGAAATTTCTCACT	2027
QY	888	GGCAATTTGAATTTGCCAAGGAATCATCCAAATATCAGGAACAGAACTCTGACTAAG	947		1921	CGGTACACCATCAAGCAGATGTTTCAAGTATGCTCTGTCTGTGGGAAATTTCTCACT	1980
DB	841	GGCAATTTGAATTTGCCAAGGAATCATCCAAATATCAGGAACAGAACTCTGACTAAG	900		2028	GGCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGGCGACGACATGGCTTACAC	2087
QY	948	GAAAAATCTTTCAGTGAACAGCTTTTTCATAGTCTGTTGATCTATGGCAAGACATTTGAC	1007		1981	GGCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGGCGACGACATGGCTTACAC	2040
DB	901	GAAAAATCTTTCAGTGAACAGCTTTTTCATAGTCTGTTGATCTATGGCAAGACATTTGAC	960		2088	CACATCAGACTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATACGA	2147
QY	1008	CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAATCAACCAAGGAGGGATGGG	1067		2041	CACATCAGACTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATACGA	2100
DB	961	CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAATCAACCAAGGAGGGATGGG	1020		2148	GGGTGGAAACGATGTCCTGAAGGAAGACCCGAATTTTCTGAAGTTGTCTGAAGTTAGAA	2207
QY	1068	CACACTGATTAACACTCTGCTGTTACCCAGGTCACATTCGCTGGTTTCAAGTTCTTACTG	1127		2101	GGGTGGAAACGATGTCCTGAAGGAAGACCCGAATTTTCTGAAGTTGTCTGAAGTTAGAA	2160
DB	1021	CACACTGATTAACACTCTGCTGTTACCCAGGTCACATTCGCTGGTTTCAAGTTCTTACTG	1080		2208	GAGTGTCTCTGCAACATTTGAGTGTCTCTGCTGATCAAGTAAACAGCAGTGGTCTCTC	2267
QY	1128	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGCTAGTGGTGA	1187				

Db 2161 GAGTGTCTCTGCAACATTGAGCTGATGTCCTCTGCAATCAAGTAACAGCAGTGGGTCTCTC 2220
Qy 2268 TCACCTTCTTCTTCTGATTCGCTGGTGAACCGGGGAGGACCTGGCCGGAGTCAATGTG 2327
Db 2221 TCACCTTCTTCTTCTGATTCGCTGGTGAACCGGGGAGGACCTGGCCGGAGTCAATGTG 2280
Qy 2328 GCAGCATTAAGAAGTGGTTCGAAATGGAATATGCTCTAAATGCAAGGTCTCTATGCTCT 2387
Db 2281 GCAGCATTAAGAAGTGGTTCGAAATGGAATATGCTCTAAATGCAAGGTCTCTATGCTCT 2340
Qy 2388 TTGTCCCAAAAGTGTGACCAATATTCCTCTCAAGGTCTCTCTTTGGAGAGATGAAAAGA 2447
Db 2341 TTGTCCCAAAAGTGTGACCAATATTCCTCTCAAGGTCTCTCTTTGGAGAGATGAAAAGA 2400
Qy 2448 AGCTTCAATACACCCATTGACCAATATGCTATGCTATGCTATGCTATGCTATGCTATGCT 2507
Db 2401 AGCTTCAATACACCCATTGACCAATATGCTATGCTATGCTATGCTATGCTATGCTATGCT 2460
Qy 2508 CATTTTCATTTCTTCCGCAATAGTAGCAGCTTTGAGGACAGCAGC 2552
Db 2461 CATTTTCATTTCTTCCGCAATAGTAGCAGCTTTGAGGACAGCAGC 2505

RESULT 6
US-09-947-199A-3
; Sequence 3, Application US/09947199A
; Patent No. 6660490
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL DARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947, 199A
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199A-3

Query Match 82.8%; Score 2505; DB 4; Length 2505;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 ATGGGAATTAATAATCTAGACCAACCCAACTTGTAAGTCAATGGAAGGAAAGTCA 107
Db 1 ATGGGAATTAATAATCTAGACCAACCCAACTTGTAAGTCAATGGAAGGAAAGTCA 60
Qy 108 AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAAA 167
Db 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAAA 120
Qy 168 GAATGACAGAACTAAGGAATATATTTGGTCTGATGAAGCTTCAAGTCAATTTA 227
Db 121 GAATGACAGAACTAAGGAATATATTTGGTCTGATGAAGCTTCAAGTCAATTTA 180
Qy 228 AATTACCGCACTGAAATGGGCTCTCTACTTCAATTTATGTTGCAATTTGTTGAGGCAAG 287
Db 181 AATTACCGCACTGAAATGGGCTCTCTACTTCAATTTATGTTGCAATTTGTTGAGGCAAG 240
Qy 288 AAATCACAATTCGAACCTTTATGTTGAAAGGGGCTCGGCCATCTCGACTCAAGAAAT 347

241 AAATCACAATTCGAACCTTTATGTTGAAAGGGCTCGGCCATCTCGACTGACAAGAAAT 300
348 GGATTTACAGCCTTGCAATTTAGCAGTTTACAGGATAATGACAGATTGATCATTCTCTG 407
301 GGATTTACAGCCTTGCAATTTAGCAGTTTACAGGATAATGACAGATTGATCATTCTCTG 360
408 CTTTACAGTGGAGCTGATATACAGCAGGTGATAGCGGTGGCTCAGTGGCTTCCATATT 467
361 CTTTACAGTGGAGCTGATATACAGCAGGTGATAGCGGTGGCTCAGTGGCTTCCATATT 420
468 GCTACAATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGCTAATGTC 527
421 GCTACAATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGCTAATGTC 480
528 AATATTCAAGATGCACTTTTTTTTCTCCATTGCAATTTGACAGCTGCTATGCAATGAA 587
481 AATATTCAAGATGCACTTTTTTTTCTCCATTGCAATTTGACAGCTGCTATGCAATGAA 540
588 CAGGTAACCTGCGCTTTTGTGAAATTTGCTGCTGATGTAATGTAAGTGGTGAAGTTGA 647
541 CAGGTAACCTGCGCTTTTGTGAAATTTGCTGCTGATGTAATGTAAGTGGTGAAGTTGA 600
648 GATAGACCCCTCCACCTAGCAGTCTGCAAAAGGATCTTGAATATTGCAAACTCTTGATG 707
601 GATAGACCCCTCCACCTAGCAGTCTGCAAAAGGATCTTGAATATTGCAAACTCTTGATG 660
708 GAAGAAAGGACGACAAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCC 767
661 GAAGAAAGGACGACAAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCC 720
768 TTCTGTCTCGATTTGGACACCATGATAGTAAGTATCTGCTGCAAGTCAATTTGGA 827
721 TTCTGTCTCGATTTGGACACCATGATAGTAAGTATCTGCTGCAAGTCAATTTGGA 780
828 GTTCAACCTCATGTTGTAATATCTATGAGATACCCCTTACACCTGCGCATGCTACAT 887
781 GTTCAACCTCATGTTGTAATATCTATGAGATACCCCTTACACCTGCGCATGCTACAT 840
888 GGCAAATTTGAAGTGTGCAAGGAAATCATCCAATATCAGGAACAGAAAGTCTGACTAAG 947
841 GGCAAATTTGAAGTGTGCAAGGAAATCATCCAATATCAGGAACAGAAAGTCTGACTAAG 900
948 GMAAACATCTTCAGTGAACAGCCTTTTCATAGTGTGCTGCTATGAGCAAGGAGCTTGAC 1007
901 GMAAACATCTTCAGTGAACAGCCTTTTCATAGTGTGCTGCTATGAGCAAGGAGCTTGAC 960
1008 CTAGTCAAAATTTCTTCTTGATCAGAAATGTATAAACATCAACCAAGGAGGATGGG 1067
961 CTAGTCAAAATTTCTTCTTGATCAGAAATGTATAAACATCAACCAAGGAGGATGGG 1020
1068 CACTGGATTACACTGCTGCTTACCAAGGATGCAATTCGCTGCTGCTGCTGCTTACTG 1127
1021 CACTGGATTACACTGCTGCTTACCAAGGATGCAATTCGCTGCTGCTGCTGCTTACTG 1080
1128 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGCTCTAGTGGTGA 1187
1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGCTCTAGTGGTGA 1140
1188 GATGAGCAGACATGTTTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 1247
1141 GATGAGCAGACATGTTTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 1200
1248 CTGAAGCATTTATAGAGACCAAGATGAAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1307
1201 CTGAAGCATTTATAGAGACCAAGATGAAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
1308 GGAGATGGCTCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1367
1261 GGAGATGGCTCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
1368 GAGAGGACAGATTTCTTCTTCAAGGCTGGAATGCTTCAATTTCCATTTCCATTTCCAT 1427
1321 GAGAGGACAGATTTCTTCTTCAAGGCTGGAATGCTTCAATTTCCATTTCCATTTCCAT 1380

1428 TCAGAAATGAGTTCATGAGATTAATGGCTCAGGTTCTTTTGGAAAGATATATAAGGA 1487
Db TCAGAAATGAGTTCATGAGATTAATGGCTCAGGTTCTTTTGGAAAGATATATAAGGA 1440
1488 CGATGAGAAATATAATAGTGGCTATAAAAGTTATCGAGCCAATACCTACTGCTCCAG 1547
Db CGATGAGAAATATAATAGTGGCTATAAAAGTTATCGAGCCAATACCTACTGCTCCAG 1500
1548 TCAGATGTGGATATGTTTGGCGAGAGGTGCCAATCTCTGCCAGCTCAATCATCCCTGC 1607
Db TCAGATGTGGATATGTTTGGCGAGAGGTGCCAATCTCTGCCAGCTCAATCATCCCTGC 1560
1608 GTAATTCAGTTTGGGTGCTTGTGTAATGATCCAGCCAGTTTGCATGTGCATCAAA 1667
Db GTAATTCAGTTTGGGTGCTTGTGTAATGATCCAGCCAGTTTGCATGTGCATCAAA 1620
1668 TACATATCAGGGGTTCTGTTCTCCCTCTTCTCATGAGCAGAGAGGATCTTGATTG 1727
Db TACATATCAGGGGTTCTGTTCTCCCTCTTCTCATGAGCAGAGAGGATCTTGATTG 1580
1728 CAGTCTAAATTAATTCAGTATGATGTTCCAAAGSCATGCGAGTACCTTCAACCTG 1787
Db CAGTCTAAATTAATTCAGTATGATGTTCCAAAGSCATGCGAGTACCTTCAACCTG 1740
1788 ACACAGCCAAATATACATCGTGACTTGAACAGTCAATATCTTCTATGAGGATGG 1847
Db ACACAGCCAAATATACATCGTGACTTGAACAGTCAATATCTTCTATGAGGATGG 1800
1848 CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAGACAAC 1907
Db CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAGACAAC 1860
1908 ATGACAAAACACCTGGAACTCGTTGGATGGCTCTGAGTGTTCACGAGTGGACT 1967
Db ATGACAAAACACCTGGAACTCGTTGGATGGCTCTGAGTGTTCACGAGTGGACT 1920
1968 CGGTACACCATCAAGAGCAGATGTTTCAAGTATGCTCTGTGTCTGTGGGAAATTCAC 2027
Db CGGTACACCATCAAGAGCAGATGTTTCAAGTATGCTCTGTGTCTGTGGGAAATTCAC 1980
2028 GGGGAAATTCATTCGCTCATCAAGCCAGCGCTGGGAGCAGACATGGCTTACCAC 2087
Db GGGGAAATTCATTCGCTCATCAAGCCAGCGCTGGGAGCAGACATGGCTTACCAC 2040
2088 CACATCAGACCTCCCATTTGGCTATTCCATTCGAGCCCATATCATCTCTGCTGATACGA 2147
Db CACATCAGACCTCCCATTTGGCTATTCCATTCGAGCCCATATCATCTCTGCTGATACGA 2100
2148 GGGTGGAAAGCATGTCCTGAAGGAGAGCCGAAATTTCTGAAGTTGTCAAGAGTTAGAA 2207
Db GGGTGGAAAGCATGTCCTGAAGGAGAGCCGAAATTTCTGAAGTTGTCAAGAGTTAGAA 2160
2208 GAGTCTCTGCAACATTTGAGTGTCTCTGATGATGATGATGATGATGATGATGATGATG 2267
Db GAGTCTCTGCAACATTTGAGTGTCTCTGATGATGATGATGATGATGATGATGATGATG 2220
2268 TCACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2327
Db TCACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2280
2328 GCAGCATTAAGAGTGGTTTGAATTTGAATATGCTCTTAAATGCAAGGTCCTATGCTGCT 2387
Db GCAGCATTAAGAGTGGTTTGAATTTGAATATGCTCTTAAATGCAAGGTCCTATGCTGCT 2340
2388 TTGTCGCAAGTGTGGACATATTTCTTCTCAAGGTCGTCTTTTGGAGGAGATGAAAAGA 2447
Db TTGTCGCAAGTGTGGACATATTTCTTCTCAAGGTCGTCTTTTGGAGGAGATGAAAAGA 2400
2448 AGTCTTCAATACACCCATGACAAATATGGCTATGATGCTGATGCTGATGCTGATGCTG 2507
Db AGTCTTCAATACACCCATGACAAATATGGCTATGATGCTGATGCTGATGCTGATGCTG 2460

2508 CATTTTCATTCTCCGAAATAGTAGCAGCTTTTGGAGACAGCAGC 2552
Db CATTTTCATTCTCCGAAATAGTAGCAGCTTTTGGAGACAGCAGC 2505

RESULT 7
US-09-458-457-7
; Sequence 7, Application US/09458457
; Patent No. 6500654
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: mnl-068CP
; CURRENT APPLICATION NUMBER: US/09/458,457
; CURRENT FILING DATE: 1999-12-10
; EARLIER APPLICATION NUMBER: 60/111,938
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/291,839
; EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3026
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(2565)
US-09-458-457-7

Query Match 66.4%; Score 2010; DB 4; Length 3026;
Best Local Similarity 80.6%; Pred. No. 0;
Matches 2425; Conservative 0; Mismatches 570; Indels 13; Gaps 6;

16 CGGCCCTTGGAGAAAGGAAACCTTATAATAATGGAAATTAATAATCTAGACCAACCC 75
Db CAGCACAGGAGAAAGCAAGAGCTTCTTTAAATGGGAATTAACAATCCAGCAACAC 88

76 AAATCTTCTACTGATGATGGAAGAAAGTCACTAGTCAATCATATTTATCACAATAGAAA 135
Db AGATCTTCTTCTGATGATGGAAGAAAGTCACTAGTCAATCATATTTATCAATAGAAA 148

136 GATTAGAGATGACCTGCAGATCAAGGAAAGAACTGACAGAACTTAAGAAATATATTG 195
Db GGCTGGAGGATTAACCTGCAGATCAAGGAAAGTCACTAGTCAATCATATTTATCAATAGAAA 208

196 GCTCTGATGAGCCCTTCAAGTCAATTTAAATTTACCGCACTGAAATGGGCTGTCTC 255
Db GCTCTGATGAGCCCTTCAAGTCAATTTAAATTTACCGCACTGAAATGGGCTGTCTC 268

256 TACTTTCATTTATGTTGCAATTTGGAGGCAAGAAATCAATATTTCAAACTCTTATGTTGA 315
Db TGCTACACCTCTGCTGTGCTGTGGGCAACAGTCAATATCCGTGCGCTTATGTTAA 328

316 AAGGGCTCCGCCATCTCGATGCAAGAAATGATTTACAGCTTGCATTTAGCAGTTT 375
Db AAGGGCTCCGTCCATCCAGACTGACGAGAAATGGGTTTCCAGCTCTGCACTGGGCGCTT 388

376 ACAAGGAAATGACAGAAATGATCACTTCTCTGCTTCAAGTGGAGCTGATATACAGCAGG 435
Db ACAAGGACAGCCCGAACTTATCACTTCTGCTGCAAGCGAGCAGATGTTGAGCAAG 448

436 TTGATACGGTGGCTCAGTCCCTCATATTTGCTCAATAGCTGCGCACTAGAGGCTG 495
Db TGGATACGGTGGCTCAGTCCCTCATATTTGCTCAATAGCTGCGCACTAGAGGCTG 508

496 CTGATGCTGTTGCAACATGAGCTAATCTCAATTTCAAGTSCAGTTTTTTTCACTC 555
Db CAGAGTGTCTGTCAACATGAGGCTCAATGTGATGTTCAAGATGCCGTCTTCTTCAACC 568

556 CATTGCATATTGACGGCTACTATGACATGAACAGGTAACCTCGCTTCTTTTGAATTTG 615

569 CACTGCACATTGCGAGCCTACTATGGGACGAGCAGGTAACCAAGTCTCTTTTGAAGTTTG 628
616 GTGCTGATGTAATGTAAGTGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAA 675
629 GTGCTGATGTAATGTAAGGCGTGAAGTTGGGACAGGCGCTCGCACCTGGCGCTCTGCAA 688
676 AAGGATTCCTTGAATATTGCAAACTCTTGAATGGAAGGACGACCAAGCAGATGTGAATG 735
689 AGGCGTCTTCAACATTTGTGAACTCTGTGTAGAAAGAGGGACCAAGCAGATGTGAAG 748
736 CTCGAAGTAATGAAGACCATGTCCCATCCCATTTCTGTCTCGATTTGGAACACCAATGATA 795
749 CTCAGGACAAATGAAGACCAAGTCCCTCTGCATCTCTGTTCTCGATTTGGAACACCAATA 808
796 TAGTTAGTATCTGCTGCAAGTGAATTTGGAAGTTCACCTCATGTGTTGTAATATCTATG 855
809 TAGTGAGCTACCTGCTCCAGAGTGACTTAGAGTCCAGGCTCAGTCATTAACATCTATG 868
856 GAGATACCCCTTACACCTGGCATGTCAATGGCAAAATTTGAAGTTGGCAAGGAAATCA 915
869 GTGACACTCTCTTTGCACCTGGCATGTCAATGGAAATTTTGAAGTTGGCAAGGAAATG 928
916 TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTCAGTGAACAGCTTTTC 975
929 TCCAGGTAACAGGAACCTGAAGTCTGACTAAGGAAACATCTTCAGCGAGACAGCTTTTC 988
976 ATAGTGTCTGTACCTATGGCAAGAGATTCACCTAGTCAAAATTTCTTTGATCAGAATG 1035
989 ACAGTGTCTGTACCTATGGCAAGACATTCAGTCACTGTGCAAAATTTCTTTGATCAGAATG 1048
1036 TCATAACATCAACCAAGGAAGGATGGGACACCTGGATACACTCTGCTGTGCTACC 1095
1049 CTGTGAACATTAACCAAGGGAAGAGATGGGACACAGGAATTCACCTCTGCTGTCTACC 1108
1096 ACCTGTACATTCGCTCGTTCAGTCTTCTACTGATATGAGAGCTGATATGAATCTAGTGG 1155
1109 ACGGCCATATCCGCTCGTTCAGTCTTCTCTGATATGAGAGCTGATATGAATCTAGTGG 1168
1156 CTTGTGATCCAGCAGGTCTAGTGGTGAAGATGAGCAGACATGTTTGAATGCTGGGCTT 1215
1169 CTTGTGATCCAGCAGGTCTAGTGGTGAAGATGAGCAGACATGTTTGAATGCTGGGCTT 1228
1216 ATGAAAAGGGCATGATGCCATTTGTCACACTCTCGAGCATATAGACACCAAGATG 1275
1229 ACAGAAAGGACATGATGCCATTTGTCACCTCTGGAAGCATACAAAGACCCCGAGAG 1288
1276 AATTGCCCTGTAATGAATATCTCAGCCTCGAGGAGATGCTCTATGTGTCTGTTCAT 1335
1289 AGCTGCCATGTAACGAATATCCAGCCTCGAGGAGATGCTCTATGTGTCTGTTCCTT 1348
1336 CACCTTTGGGGAAGATTAAGAGATGACAAAGAGAGGACAGATATTCCTCCTCCTAAGAG 1395
1349 CCCCCTTTGGGCAAGATTAAGAGATGACAAAGAGAGGAGGAGATGTTCTCCTCGAGGG 1408
1396 CTGATGTCCTTACATTTCCATCTTCAGCTCTCAGAAATGAGTTCCATGAGATTAATG 1455
1409 CTGAATACCTCCCGCTTCCATCTCCATCTCCGAAATCGAGTTCCAGAGATTAATCG 1468
1456 GCTCAGGTCTCTTTGGGAAAGTATATAAGGACGATCGAGAAATAAAATAGTGGCTATAA 1515
1469 GCTCGGCTCTCTTTGGGAAAGTCTATAAGGGCGATGCAAAATAAAATAGTGGCGATCA 1528
1516 AACGTTATCGACCAATACCTACTGCTCCAGGTCAGATGGGATATGTTTGGCGAGGG 1575
1529 AACGATACCGAGCAACACTACTGCTCCAAAGTCAGACGTTGATATGTTTGGCGAGGG 1588
1576 TGTCCATCTCTGCGAGCTCAATCATCCCTCGCTAATTCAGTTTGGGCTGTGCTTGA 1635
1589 TGTCCATCTCTGCGAGCTCAACCAACCCCTGCGTGTTCAGTTTGGGCTGTGCTGCTGG 1648
1636 ATGATCCAGGCAATTTGGCAATTTGCTCAATCAATACATATCAGGGGTTTCTCTGTTCTCCC 1695
1649 ATGACCCAGTCAAGTTTGGCAATTTGCTCACTCAGTACATTTTCAGGAGGCTCCCTGTTCTCCC 1708

1696 TCCTTCATGAGCAGAGAGGATTCCTGATTCAGTCTCTAAATTAATTTTTCAGTAGATG 1755
1709 TGCTTCATGAACAGAGAGAAATTCCTGACTTGAGTCTAAATTAATCATTTGCGTAGAGC 1768
1756 TTGCCAAAGCATGAGAGTACCTTTCAACACTGACACAGCAATTAACATCTGACTTGA 1815
1769 TTGCCAAGGSCATGGAGTACCTGACACAGCTTGACCCAGCCAATCATACCCGACCTGA 1828
1816 ACAGTCACATATTCCTCTATAGAGTGGGATGCTGTGGTGGCAGATTTTGGAGAT 1875
1829 ACAGCCACAATATTCCTCTATAGGATGGGCACTGCTGTGTGGCAGATTTTGGAGAT 1888
1876 CAAGATTTCTACAGTCTCTGGATGAAGACAAATGACAAAAAACAACCTGGGAAACCTCCGTT 1935
1889 CAAGATTTCTGAGTCCCTGGATGAAGACAAATGACAAAGCAGCCAGGAAACCTCGCT 1948
1936 GATAGGCTCTGAGGTGTTCAAGCAGTGCACTCGGTACACCATCAAGCAGATGTCTTCA 1995
1949 GGATGGCCCTGAGGTGTTTACACAGTGACGAGATACACCATCAAGGCTGATGTCTTCA 2008
1996 GCTATGCTCTGCTGTGGGAAATTCACCTGGGAAATTCATTCGCTCATCTCAAGC 2055
2009 GTTACTCCCTGCTGTGGAGTCTCTCACTGGAGAAATTCATTCGCTCATCTCAAGC 2068
2056 CAGCGGCTGGGACAGACATGGCTTACCACACATCAGACCTCCCATTTGGCTATTCCA 2115
2069 CAGCGGCTGACGACAGATATGGCGTATCACACATCAGACCGGCCCATCGGCTATTCCA 2128
2116 TTCCCAAGCCATATCATCTCTGCTATAGAGGTGGAAACCATGCTCTGAAAGAC 2175
2129 TCCCAAGCCATCTCATCCCTGCTATACGGGCTGGAAATGCATGCTCTGAAAGACGAC 2188
2176 CCGAAATTTCTGAAGTTGTCAAGTTAGAAAGTGTCTCTGCAACATTTGAGCTGATGT 2235
2189 CAGATTTCTGAGTCTGTAGCAAACTGAGAGGTGCTATGCAATGTGGAGTCTATGT 2248
2236 CTCCTGATCAAGTAAACAGAGTGGTCTCTCTACCTTCTTCTTCTGATTTGCTGG 2295
2249 CTCAGCATCAAGTAAACAGAGTGGTCTCTCTGACCTTCTCTTCTCCGATTCCTGTC 2308
2296 TGAACCGGGAGACCTGGCCGAGTCATGTGCGAGCATTAAGAAAGTCTGTTTCAATTTGG 2355
2309 TGAGCCGGGAGAGGCTTGGCCGAGCCAGTGCGAGCCTTACGGAGCCTTTTGGATTGG 2368
2356 AATATGCTCTAAATGCAAGTCTCTATGCTGTTTGTCCCAAAGTGTGACAAATATTCCT 2415
2369 AGTATGCCCTAAATGCAAGTCTCTATGCTGGGTGTTCCCAAAGTGTGGAACACACTTA 2428
2416 CTCAGGTCCTCTTTTGGAGGATGAAGAGAGTCTTCAATACACACCCATTTGACAAAT 2475
2429 ATCCGGGCTGTCTTTTGGAGGATGAATAGGAGACCCAGTATTCACCTGTTGACAAAT 2488
2476 ATGGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCAATTTCTCCGAAATAGTAGCA 2535
2489 ACGCTATGCTCTGATCCCATGAGCTGACGACCTTCACTCCCGCAAGACGACAGCA 2548
2536 GCTTTGAGGACAGAGCTGACAGCATTTGGCGTATACCTAAGAGAGATTTTTCCTCCGAA 2595
2549 ACTTTGAGGACAGCAACTGACAG-GTCTGGCATACACCTAAGGGGCTCTCCCATCAGG 2607
2596 CTGACAGCAACGATTCACACCGCAAGCTGGCTTCCAACTATATAACATTTTACTCTCAA 2655
2608 CTGACAGCATGATTTTACCCATGGCAGGCTTGTCTTCAATTAACGCGCTGCTCTG 2667
2656 AGGTCTCTTAAATTTGGGCTGTTTATCTGTCTTATTTAATTTCCCACTATTAGCAGG 2715
2668 AGGT-ITCTTTCAAAATCGTCTGTGTTTATTCTAAGCTCGTTTAAATTTCCCTTCTCAGGACAG 2726
2716 CTTTGGATTTGCTCCCTAAGGAATAATATGCAAAAGAACCAAGACAGAAATGTATGAAGA 2775
2727 GCTTTGACTCATGCC--AGCCTGAAGTGTCAAAAGAGAGATACAGAAATGTGCATGAGGA 2784

QY 2776 ATTGTTTAAATTTGTAATTTAAAAAATTTAGATCGTTACTTGGAAATGGAGCCTA 2835
DB 2785 ATTGTTCTTAGTTTGATTTTAAAGCCCTTAATTCCTGGGGCTGGGGTTCAAA---TC 2840
QY 2836 AGTCTGTGGTGACAGATAAATAATATATTTTTCCTGGGCTGAATATGTAGACTTGTGTT 2895
DB 2841 TGTGTAGATAGCTGGGTTGACCCCTATGTATTTGTAGACCAAACTGTGTGGGCTTGTGTT 2900
QY 2896 TGACAG-CTATCGGTTTATTTCTTAGACATATGTTCTATTTCTCTCATTTGTTACT 2954
DB 2901 TGAGGCTCTCTGTTGGGTTCTTAAACCAAGCTGGCTGATTTATCTCTGTTGCCCTTT 2960
QY 2955 TCTAGTGTTCACCTCTGTGATTAAAGATTTCTTGGTGAATAGAAAAAATAAAAAA 3014
DB 2961 ----GTGTGTTACTTCTGTGATTAAAGTCTCTCGTGATCTAGAAAAAATAAAAAA 3016
QY 3015 AGGGCGC 3022
DB 3017 GGGCGGC 3024

RESULT 8

US-09-947-199A-7
; Sequence 7, Application US/09947199A
; Patent No. 6660490

; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2

; CURRENT APPLICATION NUMBER: US/09/947,199A
; CURRENT FILING DATE: 2001-09-05

; PRIOR APPLICATION NUMBER: US 60/111,938
; PRIOR FILING DATE: 1998-12-11

; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1998-04-14

; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7

; LENGTH: 3026

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (61)..(2565)

US-09-947-199A-7

Query Match 66.4%; Score 2010; DB 4; Length 3026;
Best Local Similarity 80.6%; Pred. No. 0;
Matches 2425; Conservative 0; Mismatches 570; Indels 13; Gaps 6;

QY 16 CGGCCCTGGAGAAAGGAAGAACTTTATAAATGGGAATTAATACTTAGACCAACCC 75
DB 29 CAGCAGAGAGAAAGCAAGACTTCTTTAAATGGGAATTAATAATCCAGACCAACAC 88
QY 76 AAACCTGTACTGATGAATGGAAGAAAGTCAGTGAATCATATGTTATCAATAGAA 135
DB 89 AGACTGTGTTGATGAATGGAAGAAAGTTAGTGAATCTTAGCTATTATCATAGAA 148
QY 136 GATTAGAAGATCACCTGCAGATCAAGGAAAGAAAGTACAGAACTAAGGAATATATTG 195
DB 149 GGCTGGAGGATAACCTGCAGATCAAGAAATGATTTCAAGAACTAAGGCACATCTTG 208
QY 196 GCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATACCGACCTGAAATGGGCTGTCTC 255
DB 209 GCTCTGATGAAGCCTTCAGTGAAGTCAGTTTTAAATACCGACAGCGTGGCTGTCCC 268
QY 256 TACTTTCATTATGTTGCAATTTGTGGAGGCAAGAAATCACATATTCGAATCTTTATGTTGA 315
DB 269 TGCTACACCTCTGCTGTCTGTGGCGGCAAGTACATATCCGTGGCCCTTATGTTAA 328

QY 316 AAGGCTCCGCCCATCTCGACTGACAGAAATGGATTTACAGCCTTGCATTTAGCAGATTT 375
DB 329 AAGGCTCCGCTCCATCCAGACTGACGAAATGGGTTCCAGCTCTGCACCTGGCGGCTTT 388
QY 376 ACAAGGATAATGACAGAAATGATCACTTCTCTGCTTCAAGTGAGCTGATATACAGCAG 435
DB 389 ACAAGGACAGCCCGGAACCTTATCACTTCACTGTTGACAGCGGAGCAGATGTTACAGCA 448
QY 436 TTGGATAGGCTGGGCTCACTGCGCTCCATATTCCTCAATAGCTGCGCCACCTAGAGCGTG 495
DB 449 TGGGATACGCTGGGCTCACAGCCTCCATAGCTGCAATAGCTGACACCCAGAGGCTG 508
QY 496 CTGATGCTGTTGCCAATATGAGAGCTTAATGTCATATTCAGAGTCAGTTTTTTTCACTC 555
DB 509 CAGAAGTGTCTGCTACAACTAGGCGCAATGTGAATGTTCAAGATGCGCTCTCTTTCACCC 568
QY 556 CATTCATATTTGACGCTGCTAGTGGACATGAACAGGTAACTGCGCTCTTTTGAATTTG 615
DB 569 CACTGCACATTTGCAGCCTCTATGGGCACGACGAGGTAAACAGTGTCTCTTTTGAAGTTG 628
QY 616 GTGCTGATGTAATGTAAGTGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAA 675
DB 629 GTGCTGATGTCATGTAAGCGTGAAGTTGGGCAAGGCTCTGCACTGGCTCTGCAA 688
QY 676 AAGGATTCCTGAAATATTTGCAAACTCTTGTATGGAAGAGGACCAAGCAGATGTGAATG 735
DB 689 AGGCTCTTCTCAACATTTGAAAACCTCTGTTAGGAAGGAGCAAGCAGATGTGAAGC 748
QY 736 CTCGAAGATAATGAAGACCATGTCCACTCCATTTCTGTTCTCGATTTGGACACCATGATA 795
DB 749 CTCAGGACATGAAGACCATGCTCCCTCTGCACCTCTGTTCTCGATTTGGACACCAATA 808
QY 796 TAGTTAAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACTATGTTGTTAATATCTATG 855
DB 809 TAGTGAGTACTCTGCTCCAGAGTGAATTTAGAGTCCAGGCTCAAGTCATTAACATCTATG 868
QY 856 GAGATACCCCTTACACCTCGCATGTACAAATGGCAAAATTTGAAGTTGCCAAGGAATCA 915
DB 869 GTGACATCTCTTGGCCTGCGCATGCTCAATGGAATTTTGAAGTTGCCAAGGAATG 928
QY 916 TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTTCAGTGAAGAACAGCTTTTC 975
DB 929 TCCAGGTAAACAGGAACCTGAAAGTCTGACTAAGGAAACATCTTTCAGGAGACAGCTTTTC 988
QY 976 ATAGTCTTGTACTTGGCAAGACATTTGACCTAGTCAAAATTTCTTGTGATCAGAATG 1035
DB 989 ACGTGTGTTGACTATGGCAAGAACATTTGACCTGTTCAAAATTTCTTGTGATCAGAATG 1048
QY 1036 TCATAAAACATCAACCAAGGAAGGATGGGCACACTGGAATTAACCTCTGCTTGTCTACC 1095
DB 1049 CTGTGAACATTAACCAAGGAAGAGATGGGCACACAGATTTGCACTCTGCTTGTCTACC 1108
QY 1096 ACGTGTACATTCGCTGCTGCTTCACTTCTGATTAATGGAGCTGATTAATCTAGTGG 1155
DB 1109 ACGGCCATATCCGCTGCTTCACTTCTGATTAATGGTGCAGATATGAATCTGTGCG 1168
QY 1156 CTGTGTATCCAGCAGCTCTAGTGTGTAAGAAATGAGCAGACATGTTTGTGTTGGGCTT 1215
DB 1169 CTGTGTATCCAGCAGCTCTAGTGTGTAAGAAATGAGCAGACATGTTTGTGTTGGGCTT 1228
QY 1216 ATGAAAAAGGCGATGATGCTTGTGCTCACTCTCTGAAGCATTTAAGAGCCACCAAGATG 1275
DB 1229 ACAGAAAGGACATGATGCTTGTGCTTACACTCTCTGAAGCACTACAGAGACCCAGGAG 1288
QY 1276 AATTGCCCTGTAAATGATATTTCTAGCTGGAGGATGGCTCTCTATGTTGTTCCAT 1335
DB 1289 AGCTGCCATGTAAACGAATATTTCCAGCCTGGAGAGATGGCTCTCTATGTTGTTCTCT 1348
QY 1336 CACCTTTGGGGAAGATTAAAGCATGACAAAAGAGAGGACAGATATTCCTCTCTTAAGAG 1395
DB 1349 CCCCCCTTTGGGCAAGATTAAAGCATGACAAAAGAGAGGACAGATGTTCTCTCTCTGAGGG 1408
QY 1396 CTGATTTGCCTTCACATTTCCATCTTCAAGTCTCAGAAATGAGTTCCATGAGATTATTG 1455

QY	108	AGTGAATCATATGTTATCACAAATAGAAAGATTAGAAATGACCTGCGAGATCAAGGAAAA	167		1141	GATGAGCAGACATGTTTGATGTGGGCTTACGAAAGGACATGATGCCATTGTACACTC	1200	
DB	61	AGTGAATCTTACGCTATTTATCATAGAAAGGCTGAGGATTAACCTGCGAGATCAAGGAAAA	120		1248	CTGAAGCATTTATAGAGACCAAGAATGAATTCCTCCCTGTAAATGAATTTCTCAGGCTGGA	1307	
QY	168	GAACTGACAGAACTAAGAAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTTA	227		1201	CTGAAGCATTTACAGAGACCCAGAGAGAGTGCCTCAATGTAAAGAAATTTCCAGGCTGGA	1260	
DB	121	GAAATTCAGAACTAAGGCAATCTTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTTA	180		1308	GGAGATGGCTCTATGTGTCTGATCACTCACCTTGGGGAAGATTAAAAAGCATGAAAA	1367	
QY	228	AATTAACGACCTGAAAAATGGGCTGTCTCTACTTCAATTTATGTTGCAATTTGTGAGGCAAG	287		1361	GGAGATGGCTCTATGTGTCTGATCACTCACCTTGGGGAAGATTAAAAAGCATGAAAA	1320	
DB	181	AATTAACGACAGAGCGTGGGCTGTCCCTGTCTACACCTCTGCTGTCTGTGGCGCAAC	240		1368	GAGAAGCGAGATATTCCTCTCTAAGAGCTGGATTCCTTCAATTTCACTTCAGTCTC	1427	
QY	288	AAATCACATATTTGAAACTCTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAAAGAAAT	347		1321	GAGAAGCGAGATGTTCTCTCTCTGAGGGCTGAATACCTCCCGCTTCATCTCCAATC	1380	
DB	241	AAGTCACATATCCGTGGCCCTTATGTTAAAGGGCTCCGTCCATCCAGACTGACGAGAAAT	300		1428	TCAGAAATTCAGTTCCATGAGATTATGGCTCAGGTTCTTTGGGGAAGATTATAAAGGA	1487	
QY	348	GGATTTACAGCCTTGCAATTTAGCAGTTTACAGGATTAATCAGAAATTTGATCACTTCTCTG	407		1381	TCCGAAATTCAGTTCCACGAGATTATCGGCTCGGGTTCTTTGGGGAAGATTATAAAGG	1440	
DB	301	GGGTTTCCAGCTCTGCACCTGGCCGTTTACAAAGACAGCCCGGAACTTATCACTTCACTG	360		1488	CGATGCAGAAATAAATAGTGGCTATAAAACGTTATCGAGCCCAATACCTTACTGCTCCAAG	1547	
QY	408	CTTACAGTGGAGCTGATATACAGCAGTTTGGATACGGTGGCCTCACTGCCCTCCATATT	467		1441	CGATGCAGAAATAAATAGTGGCTATAAAACGTTATCGAGCCCAATACCTTACTGCTCCAAG	1500	
DB	361	TTGCAAGCGGAGCAGATGTTACAGCAAGTGGGATACGGTGGCCTTACAGCCCTTCCACATA	420		1548	TCAGATGTGATATGTTTTGCCGAGAGTGTCCATTTCTGCGAGCTCAATCATCCCTGCG	1607	
QY	468	GCTACAATAGTGGCCACCTAGAGGCTGCTGATGTGCTGTTGCAACATGGAGGTAATGTC	527		1501	TCAGACGTGATATGTTTTGCCGAGAGTGTCCATTTCTGCGAGCTCAACCACTGCTG	1560	
DB	421	GCTGCAATAGCTGGACACCCAGAGGCTGCAGAACTGCTGTACAAATGGGGCCAATGTG	480		1608	GTAATTCAGTTTGGGCTGCTTGTGAAATGATGCCAGCAGTTTGGCTTCTCACTCAA	1667	
QY	528	AAATTCAGATGAGCTGTTTTTCACTCCATGATATTCAGCGTACTATGACATGAA	587		1561	GTGTTTCAGTTTGGGCTGCTGCTGATGACCCAGTCAGTTTGGCTTCTCACTCAG	1620	
DB	481	AAATGTTCAAGATGCGCTCTTCTTCAACCCACTGACATTCAGCGCTACTATGGGACGAG	540		1668	TACATATCAGGGGTTCTGTGTTCTCCCTCTTCTATGAGCAGAGAGTTCTTCAATTG	1727	
QY	588	CAGGTAACCTCGCTCTTCTTCAAAATTTGCTGCTGATGTAATGTAAGTGGTGAAGTTGGA	647		1621	TACATTTTCAGAGGCTCCCTGTTCTCCTCTCTTCTATGAAACAGAGAAATTTCTGACTG	1680	
DB	541	CAGGTAACCTCGCTCTTCTTCAAAATTTGCTGCTGATGTAATGTAAGTGGTGAAGTTGGA	600		1728	CAGTCTAAATTAATATTGTCAGTAGATGTTGCCAAAGGATGGAGTACCTTCAACACCTG	1787	
QY	648	GATAGACCTCCACTAGCATCTGCAAAAGGATTTCTGAATATTTGCAAACTCTTGATG	707		1681	CAGTCTAAATTAATATTGTCAGTAGATGTTGCCAAAGGATGGAGTACCTTCAACACCTG	1740	
DB	601	GACAGGCTCTGCACCTGGCCTCTGCAAAAGGCTTCTTCAACATTTGTGAAACTCTGGTA	660		1788	ACACAGCAATATATACATCGTGAATTTGAAACAGTACAAATATTTCTTCTCTATGAGATGG	1847	
QY	708	GAAAGGACGACAGACAGATGTAATGCTCAAGATTAATGAAGACCATGTCACCTCCAT	767		1741	ACCCAGCAATATATACATCGTGAATTTGAAACAGTACAAATATTTCTTCTCTATGAGATGG	1800	
DB	661	GAAAGGAGGACAAAGCAGATGTAATGCTCAAGATTAATGAAGACCATGTCACCTCCAT	720		1848	CATCTGTGTGGCAGATTTTGGAGATTAAGATTTCTACAGTCTCTGATGTAAGACACAC	1907	
QY	768	TTCTGTTCTCGATTGGAACCATGATATAGTTAAGTATCTGCTGCAAACTGATTTGGA	827		1801	CATCTGTGTGGCAGATTTTGGAGATTAAGATTTCTGAGTCTCTGATGTAAGACACAC	1860	
DB	721	TTCTGTTCTCGATTGGAACCATGATATAGTTAAGTATCTGCTGCAAACTGATTTGGA	780		1908	ATGCAAAAACAACCTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTTCAAGCAGTGCCT	1967	
QY	828	GTTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAT	887		1861	ATGCAAAAACAACCTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTTCAAGCAGTGCCT	1920	
DB	781	GTCCAGGCTCAGCTCATTAAGATCTATGGTGAACACTCTTGTGACACTCTTGTGACACT	840		1968	CGGTACACCATTAAGACAGATGTTTCAAGTATGCTCTGTGTCTGTGGGAAATTTCTCCT	2027	
QY	888	GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGAAATCAGAAAGTCTGACTAAG	947		1921	AGATACACCATTAAGGCTGATGTTCTTCAAGTATGCTCTGCTGTGTGAGGCTCTCCT	1980	
DB	841	GGAAATTTGAAGTTGCCAAGGAAATTTGTCAGGTAACAGAACTGAAAGTCTGACTAAG	900		2028	GGCAAAATTTCAATTCGCTCATCTCAAGCCAGCGGCTGGCGCAGACAGATGCTTACCAAC	2087	
QY	948	GAAAAATCTTCAAGTAAACAGCTTTTCAATGAGTCTGTTGATGCTTATGCAAGAGATGAC	1007		1981	GGAGAAATTCATTCGCTCATCTCAAGCCAGCGCTGCGAGCAGCAGATATGCGGTATCAC	2040	
DB	901	GAAAAATCTTCAAGTAAACAGCTTTTCAATGAGTCTGTTGATGCTTATGCAAGAGATGAC	960		2088	CACATCAGACCTCCATTCGCTTATTCATTTCCCAAGCCCATATCATCTCTGTGTATGCA	2147	
QY	1008	CTAGTCAAAATTTCTTCTGATCAGAAATGTCATAAATCAACCAACCAAGGAGGATGGG	1067		2041	CACATCAGACCGGCTATTCGCTTATTCATCTCCCAAGCCCATCTCATCTCTGTGTATCGG	2100	
DB	961	CTGGTCAAAATTTCTTCTGATCAGAAATGTCATAAATCAACCAACCAAGGAGGATGGG	1020		2148	GGGTGGAACCATGCTCTGAAGGAGACCCGAAATTTTCTGAAGTTGTCTATGAAGTTAGAA	2207	
QY	1068	CACATGGAATTAATCTGCTGTGTAACCAAGGTCATTCGCTGTGTTTCAAGTCTTCTACTG	1127		2101	GGCTGGAATGCAATGCTCTGAAGGACCCAGAGTTCTCTGAAGTCTGTAGCAAACTGGAG	2160	
DB	1021	CACACAGGATTCATCTGCTGTGTACCAAGGCTATTCGCTGTGTTTCAAGTCTTCTACTT	1080		2208	GAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGCTGCTCAAGTAAACAGCAGTGGGTCTCTC	2267	
QY	1128	GATAATGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGCTCTAGTGTGAAAAA	1187		2161	GAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGCTGCTCAAGTAAACAGCAGTGGGTCTCTG	2220	
DB	1081	GATAATGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGCTCTAGTGTGAAAAA	1140		2268	TCACCTTCTTCTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2327	
QY	1188	GATGAGCAGACATGTTTGTGATGTGGGCTTATGAAAAAGGGCATGATGCCATTTGTACACTC	1247					

Db 2221 TCACCTTCTCTCTCCGATTCCTGCTGAGCGGGAGGCGCTGGCCGAGCCACGTG 2280
Qy 2328 GCAGCATTAAGAAAGTCTTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCTCTATGCTGCT 2387
Db 2281 GCAGCTTACGAGCGCTTTTGAGTTGAGATAGCCCTAAATGCAAGGTCTCTATGCTGGG 2340
Qy 2388 TTGTCCAAAGTGTGACAAATATTCCTCTCAAGGTCTCTTTTGGAGGAGATGAAAGA 2447
Db 2341 TGGTCCCAAAGTGTGGAACACACTCTAATCCGGGCTGTCTTTGGAGGAGATGAATAGG 2400
Qy 2448 AGTCTTCAATACACACCATTTGACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2507
Db 2401 AGACCCAGTATTCACCTGTGTGACAAATACGGCTATGTGTCTGATCCCATGAGCTGAGC 2460
Qy 2508 CATTTTCATCTTTGCCAAATAGTAGCAGCTTTTGGACAGCAGC 2552
Db 2461 CACCTTCACTCCCGCCCAAGCAGCAGCAACTTTTGGAGCAGCAAC 2505

RESULT 10
US-09-947-199A-9
; Sequence 9, Application US/09947199A
; Patent No. 6660490
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CLARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199A
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199A-9

Query Match 62.6%; Score 1892.2; DB 4; Length 2505;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 48 ATGGGAATTAATACTAGACCAACCCAACTTGCTACTGATGAATGGAAGAAAAGTC 107
Db 1 ATGGGGAATTAATACTAGACCAACCAAGCTTGCTACTGATGAATGGAAGAAAAGTT 60
Qy 108 AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 167
Db 61 AGTGAATCTTACGCTATATCATAGAAAGCTGAGGATTAACCTGCAGATCAAGAAAT 120
Qy 168 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAAGTAAGTCAATTTA 227
Db 121 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCCTTCAAGTAAGTCAATTTA 180
Qy 228 AATTACCGCACTGAAATGGGCTGCTCTACTTCTTATTTGTCATTTGTCGAGGCAAG 287
Db 181 AATTACCGCAGAGCGTGGCTGCTCCTGCTACACTCTGCTGTCTGTGTGTCGTGGCGCAAC 240
Qy 288 AAATCAATATTCGAATCTTATGTTGAAAGGCTCGGCCCATCTCGACTGACAGAAAT 347
Db 241 AAGTCAATATTCGCGCTTATGTTAAAAGGCTCCGTCATCCAGACTGACGAGAAAT 300
Qy 348 GGATTTACGCTTGCATTTAGCAGTTTACAGGATTAATGCGAATGATCACTTCTCTG 407

301 GGGTTTCCAGCTCTGCACCTGGCCGTTTACAAAGACAGCCCGAACTTATCACTTCACTG 360
408 CTTTCAGAGTGGAGCTCATATACAGAGGTTCGATACGGTGGCTCAGTCCCTCCATATT 467
361 TTGCAACAGCGGACAGATGTTCAAGAGTGGATACGGTGGCTCACAGCCCTCCACATA 420
468 GCTCAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTTTGCCAACATGGAGCTAATGTC 527
421 GCTGCAATAGCTGGACACCCAGAGGCTGAGAGGTGCTGTACAACTGGGGCCAAATGTG 480
528 AATATTCAAGATGCACTTTTTCATCCATTCGATTTGTCAGCTGCTTATGCGACATGAA 587
481 AATGTTCAAGATGCGCTTTCTTCCACCCACTGCAATTCGAGCTCTATATGGGACGAG 540
588 CAGGTAACCTGCGCTTTCTTTGAAATTTGCTGCTGATGTAATGTAAGTGGTGAAGTGA 647
541 CAGGTAACCAAGTGTCTTTTGAAGTTTGGTGTGATGTCAATGTAAAGCGGTGAAGTTGG 600
548 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATCTTGAATATTGCAAAACTCTTGATG 707
601 GACAGGCTCTGCACCTGGCTCTGCAAAAGGCTTCTTCAACATTTGAAACTCTCTGTA 660
708 GAAAGAGGACAGCAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCATCCAT 767
661 GAAGAGGGAGCAAGCAGATGTGAACGCTCAGGACATGAAGACCATGCTCCTCTGCAC 720
768 TTTCTGTTCTGATTTGGACACCAATGATTAAGTATCTGCTGCAAGATGATTGGA 827
721 TTTCTGTTCTGATTTGGACACCAATATAGTAGTACCTCTCCAGAGTACTTAGAG 780
828 GTTCAACCTCATGTTGTAATATCTATGGAGATACCCCTTACACTGACCTGGCATGTACAT 887
781 GTCCAGCTCAGCTCAATTAACATCTATGAGACATCTCTTTCGACCTGGCATGCTACAT 840
888 GGCAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947
841 GGAATTTTGAAGTTGCCAAGGAAATTTGCCAGGTAAACAGGAATGAAAGTCTGACTAAG 900
948 GAAACATCTTTCAGTGAACACAGCTTTTCATAGTGTGCTGATGCTATGGCAAGAGCATGAC 1007
901 GAAACATCTTTCAGCGAGACAGCTTTTCACAGTGTGTTGACCTATGGCAAGAACTTGAC 960
1008 CTAGTCAAAATTTCTTCTGATCAGAAATGTCAATAACATCAACCAACAGGAAGGATGG 1067
961 CTGTCAAAATTTCTTCTGATCAGAAATGTGTGAACATTAACCCAGGAGAAAGATGG 1020
1068 CACACTGGATTAACATCTGCTGCTACAGGTCACATTCGCTGGTTCAGTTCTTACTG 1127
1021 CACACAGGATTTGACATCTGCTGCTACACGGCCATATCCGCTGGTTCAGTTCTTACTT 1080
1128 GATAATGGAGCTCATATGAATCTAGTGGCTTGTGATCCCGCAGAGCTCTAGTGGTGAATA 1187
1081 GATAATGGTGCAGATATGAATCTTGTGCTTGTGATCCCGCAGAGCTCTAGTGGTGAATA 1140
1188 GATGAGCAGACATGTTTGTGATGTTGGCTTATGAAAAGGCGCATGATGCCATTTGACATC 1247
1141 GATGAGCAGACATGTTTGTGATGTTGGCTTACGAGAAAGGACATGATGCCATTTGACATC 1200
1248 CTGAAGCATTTAAGAGACACAGATGAATTCCTGCTGAATGAATATTTCTCAGGCTGA 1307
1201 CTGAAGCATTTAAGAGACACAGGAGCTGCCATGTAAGCAATATTTCCAGGCTGGA 1260
1308 GAGATGGCTCTTATGCTGTGTTCCATCACCTTGGGAAAGATTAAAGCATGACAAAA 1367
1261 GGAGATGGCTCTTATGCTGTGTTCTTCCCTTGGGCAAGATTAAAGCATGACAAAA 1320
1368 GAGAGGCGAGATTTCTCTCTTAAGAGCTGGAATTCCTTACATTTCCATTTCCAGTCTC 1427
1321 GAGAGGCGAGATTTCTCTCTGAGGCTGAACTACCTCCGCTTCCATCTCCAACTC 1380
1428 TCAGAAATTCAGTTTCCATGAGATTAATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGA 1487
1381 TCCGAATTCAGTTTCCACGAGATTAATTCGGCTCGGGTCTCTTTGGGAAAGTCTATAAGGG 1440

```
QY 1488 CGATCGAATAAATAAGTGGCTATATAACGTTATCGAGCCAAATACCTACTGCTCCAG 1547
Db 1441 CGATCGAATAAATAAGTGGCTATATAACGTTATCGAGCCAAATACCTACTGCTCCAG 1500
QY 1548 TCAGATGTGGATATGTTTGGCGAGAGGTGCCATCTCTGCCAGCTCAATCATCCTGCG 1607
Db 1501 TCAGACGTGGATATGTTTGGCGAGAGGTGCCATCTCTGCCAGCTCAATCATCCTGCG 1560
QY 1608 GTAATTCAGTTTGTGGGTGCTTGTGTAATGATCCAGCCAGTTTGGCCATGTGCACTCAA 1667
Db 1561 GTGGTTCAAGTTTGTGGGTGCTTGTGTAATGATCCAGCCAGTTTGGCCATGTGCACTCAA 1620
QY 1668 TACATATCAGGGGTTCTCTGTTCTCCCTCTTATGATGAGAGAGGATCTTGATTG 1727
Db 1621 TACATTTAGGAGGCTCCCTGTTCTCCCTGTTCTATGAACAGAGAGAAATCTTGACTTG 1680
QY 1728 CAGTCTAAATTAATTTGCAAGTAGATGTTGCCAAGGCATGGAGTACCTTCACAACTTG 1787
Db 1681 CAGTCTAAATTAATTTGCGGTAGACGTTGCCAAGGCATGGAGTACCTTCACAGCTTG 1740
QY 1789 ACACAGCCAAATATACATCGTGACTTGAAAGTGCACAAATATCTCTCTATGAGGATGGG 1847
Db 1741 ACCCAGCCAAATCATACCGCGACCTGAACAGCCACAAATATCTGCTCTATGAGGATGGC 1800
QY 1848 CATGCTGTGGTGGCAGATTTTGGAGAATCAAGATTTCTACAGTCTCTGGATGAAGACAA 1907
Db 1801 CATGCTGTGGTGGCAGATTTTGGAGAATCAAGATTTCTGAGTCCCTGGATGAAGACAA 1860
QY 1908 ATGACAAACAACTGGGAACTCCGTTGGATGCTCTGAGGTGTTTCACGAGTGCAC 1967
Db 1861 ATGACAAACGACCCAGGAACTCCGCTGGATGCTGAGTGGTTCACACAGTGCACG 1920
QY 1968 CGGTACACATCAGAGCAGATGCTTCAGCTATGCTCTGTGCTGTGGGAAATTCAC 2027
Db 1921 AGATACACATCAGAGCAGATGCTTCAGTTCCTCTGCTGTGGAGTCTCTCACT 1980
QY 2028 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGGCGAGCAGACATGGCTTACCAC 2087
Db 1981 GGAGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGGCGAGCAGATATGGCGTATCAC 2040
QY 2088 CACATCAGACCTCCCATTTGGCTATTCATTTCCGAGGCCATATCATCTCTGCTGATACGA 2147
Db 2041 CACATCAGACCGCCCATCGGCTATTCATTCGCCAAGGCCATCTCATCCCTGCTGATACGG 2100
QY 2148 GGGTGGAAAGCATGCTCGTGAAGGAGACCCGAAATTTCTGAAGTTGTCATGAAGTTAGAA 2207
Db 2101 GGTGGAATGCATGCTGAAGGACGACAGAGTTCTGAGTCTGTTACGNAACTGGAG 2160
QY 2208 GAGTGTCTGCAACATTTGAGTCTCTCTGATCAAGTAAAGCAGTGGGTCTCTC 2267
Db 2161 GAGTGCCTATGCAATGTGGAGCTCATGTCTCCAGCATCAAGTAAAGCAGTGGGTCTCTC 2220
QY 2268 TCACCTCTCTCTCTCTGATTTGCTGTGAACCGGAGAGCTGCGCGAGTCAATGTG 2327
Db 2221 TCACCTCTCTCTCTCTGATTTGCTGTGAACCGGAGAGCTGCGCGAGAGTCAATGTG 2280
QY 2328 GCAGCATTAAGAGTCGTTTGCATTTGGAATATGCTCTAAATGCAAGGTCTTATGTGCT 2387
Db 2281 GCAGCCTTACGAGCGGTTTGTGAGTGTGAGTATGCCCTAAATGCAAGGTCTTATGTG 2340
QY 2388 TTGTCGCAAGTCTGGACATATTTCTCTCAAGGTCTGTTTGGAGGAGTGAAGA 2447
Db 2341 TGGTCCCAAGGTTTGGAAACACATCTTAATCCGGGCTGCTTCTTGGAGGAGTGAATAGG 2400
QY 2448 AGTCTTCAATACACACCCATTTGACAAATATGAGTATGATCCGATCCATGAGCTCAATG 2507
Db 2401 AGCACCCAGTATTTCAACTGTTGACAAATACGGCTATGTGCTGATCCATGAGCCTGACG 2460
QY 2508 CATTTTCAATTTGCGGAAATAGTAGCAGCTTTGAGGACAGCAGC 2552
Db 2461 CACCTTCACTCCCGCCCAAGACGACAGCAACTTTTGGAGCAGCAGC 2505
```

```
RESULT 11
US-09-833-381-1151
; Sequence 1151, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIORITY FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1151
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(740)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1151

Query Match      14.1%; Score 427.6; DB 4; Length 740;
Best Local Similarity 78.4%; Pred. No. 7.4e-107;
Matches 519; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

QY 16 CCGCCCTGGAGAAAGGAGAACTTATATAAATGGGAAATTTATAATCTACCAACCC 75
Db 41 CAGCACAGGAGAAAGCAAGACTTCTTTAAATGGGAAATTAACAATCCAGCAACAC 100
QY 76 AAATCTGTACTGTGATGAATGAAGAAAGTCAAGTCAATCATATGTTATCAAAATAGAAA 135
Db 101 AGACTTGTCTGATGAATGAAGAAAGTGTAGTGAATCTTACGCTATTATCATAGAAA 160
QY 136 GATTAGAAATGACCTGCAGATCAAGGAAAGAAAGTGCAGAACTTAAGGAATATTTG 195
Db 161 GGCTGGAGGATAACCTGCAGATCAAAAGAAATGAATTTCAAGAACTTAAGGCAATCTTTG 220
QY 196 GCTCTGATGAAGCTTCAGTAAAGTCAATTTAAATTTACCGCACTGAAATGGCTGTCTC 255
Db 221 GCTCTGATGAAGCTTCAGTAAAGTCAATTTAAATTTACCGCACTGAAATGGCTGTCTC 280
QY 256 TACTTCAATTTATGTTGCAATTTGTGGAGGCAAGAAATCAATATTCGAATCTTATGTTGA 315
Db 281 TGCTACACCTCTGCTGTGCTGTGGCGCAACAAGTCAATATCCGTGCCCTTATGTTAA 340
QY 316 AAGGCTCCGCCATCTCGACTCAAGAAATGGATTTACAGCTTGCATTTAGCAGTTT 375
Db 341 AAGGCTCCGCCATCTCGACTCAAGAAATGGATTTACAGCTTGCATTTAGCAGTTT 400
QY 376 ACAGGATAATGAGAAATTCATCTTCTCTGCTTCACT-AGTGGAGCTGATATACAGCAG 434
Db 401 ACAGGACAGCCGGAATTCATCTTCACTGTGCAAGCGGAGCAGATGTTTCAGCAA 460
QY 435 GTTGGATACGGTGGCTCTACCTCCCTCCATATTTGCTAATAGCTGGCCACCTAGAGCT 494
Db 461 GTGGGATACGGTGGCTCTACAGCCCTCCCATAGCTGCAATAGCTGGACACCCAGAGCT 520
QY 495 GCTGATGCTGTGTTGCAACATGGAGTAAATGTCAATATTCAAGATGTCAGTTTCTTCACT 554
Db 521 GCANAAGTGTGNTGCAACATGGGGCAACGTGAATGTTCAAGATGCCGTCTTCTTACC 580
QY 555 CCATTGCAATTTGAGCGTACTATGGACATGAACAGGTAACCTGCGCTTCTTTGAAATTT 614
Db 581 CCATGTCACATTTGAGCGTACTATGGGACGAGCAGGTAACCAAGTGTCTCTTTGAAGTT 640
QY 615 GGTGCTCATGTAATTTGAGTGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCA 674
Db 641 GGNGGTGATGTCATGTAAGCGGTNAAGTTGGNACAGGCTCCGCCCTCGCCNCGNA 700
```

QY 675 AA 676
||
Db 701 AA 702

RESULT 12

US-09-833-381-1149
; Sequence 1149, Application US/09833381
; Patent No. 6672186

GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1149
; TYPE: DNA

; LENGTH: 616
; ORGANISM: Homo sapiens

US-09-833-381-1149

Query Match 14.1%; Score 427.4; DB 4; Length 616;

Best Local Similarity 81.8%; Pred. No. 7.6e-107;

Matches 505; Conservative 0; Mismatches 111; Indels 1; Gaps 1;

QY 338 GACAGAAATGGATTTACAGCTTGTAGAGTTTACAGGATTAACAGGATAATGCAGAAATTGAT 397
||
Db 1 GACAGAAATGGGTTTCCAGCTCTGCACCTGGCGGTTTACAGGACAGCCCGGAATTAT 60

QY 398 CACTTCTCTGCTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCACTGC 457
||
Db 61 CACTTCACTGTGACAGCGGAGCAGATGTTTACAGGATGGGATACGGTGGCTCACAGC 120

QY 458 CCTCATATTGCTACATAGCTGGCCACCTAGAGGCTGTGATGCTGCTTCCACATGG 517
||
Db 121 CCTCC-CATAGCTGCAATAGCTGACACCCAGAGGCTGAGAAAGTCTCTGCAACATGG 179

QY 518 AGCTAAATGTCAAATATCAAGATGACGATTTTTCCTCCATTCATTCATTCAGTGTGAGCGTACTA 577
||
Db 180 GGCCACAGTGAATGTTCAAGATGCGCTTCTTCCACCCACTGCAATTCGAGCCTACTA 239

QY 578 TGGACATGAACAGATTAAGTCTCTTTTGAATTTGGTGTGATGTAATGTAAGTG 637
||
Db 240 TGGGCACGACGAGTAAACCAAGTGTCTTTTGAAGTTTGGTGTGATGTCATGTAAGCGG 299

QY 638 TGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTCRAA 697
||
Db 300 TGAAGTTGGGACAGGCTCTGCACCTGGCTCTGCAAGGGGTTCTTCAACATTTGAA 359

QY 698 ACTCTTGATGAAGAAGGACGCAAGCAGATGTAATGCTTCAAGATTAATGAAGACCATG 757
||
Db 360 ACTCCTGTTGAAGAAGGAGCAAGCAGATGTAACGCTTCAGCAATGAAGACCACTG 419

QY 758 CCCACTCCATTTCTGTTCTGATTTGGACACCATGATATAGTTAACTATCTGCTGCAAG 817
||
Db 420 CCCCTGCACTTCTGTTCTGATTTGGACACCAATATAGTAGCTACTGCTCCAGAG 479

QY 818 TGATTTGGAAGTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGCG 877
||
Db 480 TGACTTAGAGGTCAGCGCTCAGCGTCATTAAACATCTATGGTGAACACTCTTTGCACTGCG 539

QY 878 ATGCTACAAATGGCAAAATTTGAAGTTGCAAGGAAATCATCCAAATATCAGGAACAGAA 937
||
Db 540 ATGCTACAAATGGAAATTTGAAGTTGCAAGGAAATTTGTCAGGTAACAGGAACCTGAAG 999

QY 938 TCTGACTAAGGAAACA 954
||
Db 600 TCTGACTAAGGAAACA 616

RESULT 13

US-09-833-381-1150
; Sequence 1150, Application US/09833381
; Patent No. 6672186

GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1150
; LENGTH: 736

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)..(736)

; OTHER INFORMATION: n = A,T,C or G

US-09-833-381-1150

Query Match 13.7%; Score 414.6; DB 4; Length 736;

Best Local Similarity 81.2%; Pred. No. 2.6e-103;

Matches 480; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 312 TTGAAGGGCTCCGCCATCTCGACTGACAAGAAATGGATTTACAGCCTTGCAATTTAGCA 371
||
Db 1 TTAAAGGGCTCCGTCATCCAGACTGACGAGAAATGGGTTCCAGCTCTGCACCTGGCC 60

QY 372 GTTTACAAGGATTAATGCAGAAATGATCATCTTCTCTCTCAAGTGGAGCTGATATACAG 431
||
Db 61 GTTTACAAGGACAGCCCGAACTTATCACTTCACTGTTGACAGCGGAGCAGATGTTCA 120

QY 432 CAGTTTCGATACCGTGGCCTCACTGCGCTCCATTTGCTTACATAGCTGCGCACCTAGAG 491
||
Db 121 CAGTTGGATACCGTGGCCTCACAGCCTCCATATAGCTGCAATAGCTGACACCCAGAG 180

QY 492 GCTGCTGATGCTGTTTGGCAACATGGAGCTTAATGTCAATATTTCAAGATCAGTTTTTTC 551
||
Db 181 GCTGCAAGAGTGTCTTACAACATGGGCAATGTGAATGTTCAAGATCCCGTCTTCTTC 240

QY 552 ACTCCATTCATATTCAGCGTACTATGACATGAACAGTAACTGCGCTTCTTTGAAA 611
||
Db 241 ACCCACTGCACATTCGAGCCTACTATGGGACGACGAGTAAACAGTGTCTCTTTGAAG 300

QY 612 TTTGGTCTCATGTAAATGTAAAGTGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCT 671
||
Db 301 TTTGGTCTCATGTCAATGTAAAGGTTGAAGTTGGGACAGGCTCTGCACCTGGCCTCT 360

QY 672 GCAAAAGGATTTCTGAATATTTGCAAACTCTTTGATGGAAAGGAGCAAGCAGATGTG 731
||
Db 361 GCAAAAGGCTTTCTCAACATTTGTAAACTCTCTGTTAGAAAGGAGCAAGCAGATGTG 420

QY 732 AATGCTCAAGATTAATGAAGACCATGTCCTCACTCCATTTCTGTTCTGATTTGGACACCAT 791
||
Db 421 AACCTCAGGACAAATGAAGACCACTGTCCTGTCGACTTCTGTTCTGATTTGGACACCA 480

QY 792 GATATAGTTAAATGATCTGCTGCAAGTGTATTTGGAAGTTCAACCTCATGTTGTTAATC 851
||
Db 481 AATATAGTAGCTACCTGCTCCAGAGTGACTTANAGGTTCCAGCTCCAGCTCATTAACATC 540

QY 852 TATGGAGATACCCCTTACACCTGGCATGCTACATGGCAATTTGAAGTT 902
||
Db 541 TATGGTGACACTCCTTTGACCTGGCATGCTCAATGGAAATTTTGAANTT 591

RESULT 14

US-09-833-381-1147/c

; Sequence 1147, Application US/09833381

Patent No. 6672186
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
FILE REFERENCE: 5800-119
CURRENT APPLICATION NUMBER: US/09/833,381
CURRENT FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/516,448
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 2050
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1147
LENGTH: 304
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(304)
OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1147

Query Match 6.1%; Score 184.2; DB 4; Length 304;
Best Local Similarity 79.6%; Pred. No. 1.4e-40;
Matches 242; Conservative 0; Mismatches 59; Indels 3; Gaps 2;

QY 313 TGAAGGGCTCGCCCATCTCGACTGACAAGAAATGGATTTACAGCC--TTGCATTTAGC 370
DB |||||
304 TAAAGGGCTCGCTCATCCAGACTGACGAGAAATGGTTTCCAGATTTTGACCTGCC 245
QY 371 AGTTTACAAGGATAATGAGAATTGATCAC-TTCTCTGCTTCAAGTGGAGCTGATATAC 429
DB |||||
244 GGTTTACAAGGACAGCCCGGAATTTATCACTTTCACTGTTGCACAGCGGACAGATGTTT 185
QY 430 ACAGAGTTGGATACCGTGGCTCACTGCCCTCCATATGCTACATAGCTGGCCACTAG 489
DB |||||
184 AGCAAGTGGATACCGTGGCTCACTGCCCTCCATATGCTACATAGCTGGCCACTAG 125
QY 490 AGGCTGCTGATGCTGTTGTCACAAATGGAGCTAATGTCAATATTTCAAGATGCAGTTTTT 549
DB |||||
124 AGGCTGCAAGAGTCTGATGCAACATGGGCCAACGTGAATGTTCAAGATGCCGTCTTCT 65
QY 550 TCACTCCATTCATATTCAGGCTACTATGACATGACAGGTAACCTGCGCTTCTTTGA 609
DB |||||
64 TCACCCCACTGCATTCGACGCTACTATGGGCACGACGAGTAACCAAGTGTCTCTTTGA 5
QY 610 AATT 613
DB 4 AGTT 1

RESULT 15
US-09-833-381-1148/c
Sequence 1148, Application US/09833381
Patent No. 6672186
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
FILE REFERENCE: 5800-119
CURRENT APPLICATION NUMBER: US/09/833,381
CURRENT FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/516,448
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 2050
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1148
LENGTH: 186
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(186)
OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1148

Query Match 2.8%; Score 84.4; DB 4; Length 186;
Best Local Similarity 69.1%; Pred. No. 1.9e-13;
Matches 112; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 452 CACTGCCCTCCATATTCCTACAATAGCTGGCCACCTAGAGGCTGCTGTGCTGTGCA 511
DB |||||
162 CCCTTCCAANATAGTCCCAATAGTTGNACACCCCAANAGGTGCAAAAGTGTGTGCA 103
QY 512 ACATGGAGCTAATGCTCAATATTCAGATGCAAGTTTTCCTCCATTCGATATTCAGC 571
DB |||||
102 ACATGGGCGCAATGTGAATGTTCAAGATGCCGCTCTTTCACCCCACTGCACATTCGAGC 43
QY 572 GTACTATGGACATGAACAGGTAACCTGCCCTTCTTTTGAATT 613
DB |||||
42 CTACTATGGGACGAGGAGGTAAACCAGTGTCTCTTTTGAAGTT 1

Search completed: September 6, 2004, 07:42:49
Job time : 152.506 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2004, 05:03:46 ; Search time 915.158 Seconds
(without alignments)
16430.147 Million cell updates/sec

Title: US-10-626-173-1

Perfect score: 3025

Sequence: 1 GTGACCCAGCGTCGGCCG.....aaaaaaagggcgccgc 3025

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3267054 seqs, 2485319735 residues

Total number of hits satisfying chosen parameters: 6534108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	3025	100.0	3025	9	US-09-947-199-1
2	3025	100.0	3025	17	US-10-626-173-1
3	2505	82.8	2505	9	US-09-947-199-3
4	2505	82.8	2505	17	US-10-626-173-3
5	2010	66.4	3026	9	US-09-947-199-7
6	2010	66.4	3026	17	US-10-626-173-7
7	1892.2	62.6	2505	9	US-09-947-199-9
8	1892.2	62.6	2505	17	US-10-626-173-9
9	427.6	14.1	740	9	US-09-833-381-1151
10	427.4	14.1	616	9	US-09-833-381-1149
11	414.6	13.7	736	9	US-09-833-381-1150
12	184.2	6.1	304	9	US-09-833-381-1147
13	109.6	3.6	1953	13	US-10-424-599-53404
14	107.4	3.6	3238	13	US-10-424-599-129156

15 100.2 3.3 2207 13 US-10-425-114-22955 Sequence 22955, A
16 100.2 3.3 2327 17 US-10-437-963-54440 Sequence 54440, A
17 98 3.2 1662 9 US-09-938-842A-1014 Sequence 1014, Ap
18 98 3.2 1662 11 US-09-938-842A-1014 Sequence 1014, Ap
19 97 3.2 1572 13 US-10-425-114-33402 Sequence 33402, A
20 91.4 3.0 1369 13 US-10-425-114-30542 Sequence 30542, A
21 91.4 3.0 1545 13 US-10-425-114-21691 Sequence 21691, A
22 91.4 3.0 1614 13 US-10-425-114-13674 Sequence 13674, A
23 91.4 3.0 3323 16 US-10-104-047-313 Sequence 313, App
24 89.4 3.0 972 13 US-10-425-114-6886 Sequence 6886, Ap
25 88.6 2.9 1146 17 US-10-437-963-59479 Sequence 59479, A
26 88.6 2.9 1461 13 US-10-425-114-4043 Sequence 4043, Ap
27 87.8 2.9 1219 13 US-10-424-599-130828 Sequence 130828, A
28 87.8 2.9 3020 17 US-10-437-963-65894 Sequence 65894, A
29 87.4 2.9 1299 9 US-09-908-711-12 Sequence 12, Appl
30 87.2 2.9 2919 17 US-10-437-963-70898 Sequence 70898, A
31 85.6 2.8 5338 13 US-10-342-887-63 Sequence 63, Appl
32 85.6 2.8 5338 13 US-10-172-118-63 Sequence 63, Appl
33 85 2.8 1942 13 US-10-425-114-28428 Sequence 28428, A
34 85 2.8 1997 13 US-10-425-114-30767 Sequence 30767, A
35 84.4 2.8 186 9 US-09-833-381-1148 Sequence 1148, Ap
36 84.2 2.8 5175 9 US-09-964-899-42 Sequence 42, Appl
37 84 2.8 799 13 US-10-425-114-30676 Sequence 30676, A
38 84 2.8 14770 10 US-09-873-367C-230 Sequence 230, App
39 84 2.8 14770 12 US-09-968-007A-124 Sequence 124, App
40 84 2.8 14770 12 US-09-968-007A-405 Sequence 405, App
41 84 2.8 14770 13 US-10-342-887-1848 Sequence 1848, Ap
42 84 2.8 14770 13 US-10-172-118-1848 Sequence 1848, Ap
43 83.6 2.8 2473 17 US-10-437-963-92781 Sequence 92781, A
44 83.4 2.8 1991 17 US-10-437-963-93604 Sequence 93604, A
45 82.6 2.7 4213 16 US-10-108-260A-794 Sequence 794, App

ALIGNMENTS

RESULT 1

US-09-947-199-1
; Sequence 1, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-068CE2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2552)
US-09-947-199-1

Query Match 100.0%; Score 3025; DB 9; Length 3025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTGACCCAGCGTCGGCCCTGGAGAAAGGAAGAACTTATAATAATGGGAATTATA 60
Db 1 GTGACCCAGCGTCGGCCCTGGAGAAAGGAAGAACTTATAATAATGGGAATTATA 60
Qy 61 AATCTAGACCAACCCAACTTGTACTGTAGTGAAGAAAAAGTCAATCATATG 120

QY 2281 CTTCTGATTGCTGGTGAACCGGGGAGGAGCTGCGCGAGTCAATGCGAGCAATTAAGAA 2340
Db CTTCTGATTGCTGGTGAACCGGGGAGGAGCTGCGCGAGTCAATGCGAGCAATTAAGAA 2340
QY 2341 GTCGTTTCGAATTCGAATATGCTCTAAATCAAGGTCCTATGCTGCTTGTGTCCTCAAGTG 2400
Db GTCGTTTCGAATTCGAATATGCTCTAAATCAAGGTCCTATGCTGCTTGTGTCCTCAAGTG 2400
QY 2401 CTGGACAATATTCCTCTCAAGGTCCTGCTTGTGGAGGAGATGAAAAGAGTCTTCAATACA 2460
Db CTGGACAATATTCCTCTCAAGGTCCTGCTTGTGGAGGAGATGAAAAGAGTCTTCAATACA 2460
QY 2461 CACCCATTGCAATATGCTGATGATCCGATCCCATGAGCTCAATGCAATTTTCAATCTT 2520
Db CACCCATTGCAATATGCTGATGATCCGATCCCATGAGCTCAATGCAATTTTCAATCTT 2520
QY 2521 GCCGAATAGTAGCAGCTTTGAGGACAGCAGCTGACAGCAATTCGGGCTATACCTAAGGAG 2580
Db GCCGAATAGTAGCAGCTTTGAGGACAGCAGCTGACAGCAATTCGGGCTATACCTAAGGAG 2580
QY 2581 AGTTTTTCCCGAATGACAGCAAGATTCGAACCAACGCAAGCTGGCTTCCAACTATA 2640
Db AGTTTTTCCCGAATGACAGCAAGATTCGAACCAACGCAAGCTGGCTTCCAACTATA 2640
QY 2641 ACATTTTACTCTCAAGGTCCTCTTAAATGGGCTGTTTTTACTTGTCTCTATTTAATTC 2700
Db ACATTTTACTCTCAAGGTCCTCTTAAATGGGCTGTTTTTACTTGTCTCTATTTAATTC 2700
QY 2701 CCCACTATTAGCAGCTTTGGATTTGCGCTAAGGAATATATGCAAAAGAACCAAGACA 2760
Db CCCACTATTAGCAGCTTTGGATTTGCGCTAAGGAATATATGCAAAAGAACCAAGACA 2760
QY 2761 GAATGTATATGAAGATTTGTTTTAATTTGTAATTTAAATTTAAATTTAGATCGTTACT 2820
Db GAATGTATATGAAGATTTGTTTTAATTTGTAATTTAAATTTAAATTTAGATCGTTACT 2820
QY 2821 TGGAAATGGAGCCCTAAGTCTGCTGGTGGACAGATAATATGTTTCTCGGCTGAAAT 2880
Db TGGAAATGGAGCCCTAAGTCTGCTGGTGGACAGATAATATGTTTCTCGGCTGAAAT 2880
QY 2881 ATGTAGACTTGTGTTGACAGCTATGGGTTTATTTCTTAGACATTTGTTTCTTTT 2940
Db ATGTAGACTTGTGTTGACAGCTATGGGTTTATTTCTTAGACATTTGTTTCTTTT 2940
QY 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGCGATTAAAGATTTCTTGGTGAATAGAAA 3000
Db CTCATTATGTTACTTCTAGTGTTCACCTCTGCGATTAAAGATTTCTTGGTGAATAGAAA 3000
QY 3001 AAAAAAAAAAAAAAGGGCGCGC 3025
Db AAAAAAAAAAAAAAGGGCGCGC 3025

RESULT 2

US-10-626-173-1
; Sequence 1, Application US/10626173
; Publication No. US20040110232A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MN-068CF2
; CURRENT APPLICATION NUMBER: US/10/626,173
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: US/09/947,199A
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/231,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2552)
US-10-626-173-1

Query Match 100.0%; Score 3025; DB 17; Length 3025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTCACCCACGCGTCCGCGCTGGAGAAAGAAAGAACTTATTAATAATGGGAATATATA 60
Db 1 GTCACCCACGCGTCCGCGCTGGAGAAAGAAAGAACTTATTAATAATGGGAATATATA 60
QY 61 AATCTAGACCAACCCAAACTTTGATGTAATGGAAGAAAGAAAGTCACTGAATCATATG 120
Db 61 AATCTAGACCAACCCAAACTTTGATGTAATGGAAGAAAGAAAGTCACTGAATCATATG 120
QY 121 TTATCACAATAGAAAGATTAGAAAGATGACCTGCAGATCAAGGAAAGAAAGAACTGACAGAAC 180
Db 121 TTATCACAATAGAAAGATTAGAAAGATGACCTGCAGATCAAGGAAAGAAAGAACTGACAGAAC 180
QY 181 TAAGGAATATTTGGCTCTGATGAGACCTTCAGTAAAGTCAATTTAAATTTACCGCACTG 240
Db 181 TAAGGAATATTTGGCTCTGATGAGACCTTCAGTAAAGTCAATTTAAATTTACCGCACTG 240
QY 241 AAAATGGGCTGTCTCTACTCTTCAATTTATGTTGCAATTTTGGAGGCAAGAAATCACATATTC 300
Db 241 AAAATGGGCTGTCTCTACTCTTCAATTTATGTTGCAATTTTGGAGGCAAGAAATCACATATTC 300
QY 301 GAACCTCTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAATGGATTTACAGCT 360
Db 301 GAACCTCTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAATGGATTTACAGCT 360
QY 361 TGCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTGCTTACAGTGGAG 420
Db 361 TGCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTGCTTACAGTGGAG 420
QY 421 CTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATTTGCTACAAATAGCTG 480
Db 421 CTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATTTGCTACAAATAGCTG 480
QY 481 GCCACCTAGAGGCTGCTGATGCTGCTGCAACATGGAGCTAATGTCATATTTCAAGATG 540
Db 481 GCCACCTAGAGGCTGCTGATGCTGCTGCAACATGGAGCTAATGTCATATTTCAAGATG 540
QY 541 CAGTTTTTTTCACTCCATTTGCATATTTGCAGCGTACTATGGACATGAACAGGTAACCTCGCC 600
Db 541 CAGTTTTTTTCACTCCATTTGCATATTTGCAGCGTACTATGGACATGAACAGGTAACCTCGCC 600
QY 601 TTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTGGAGATAGACCCCTCC 660
Db 601 TTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTGGAGATAGACCCCTCC 660
QY 661 ACCTAGCATCTGCAAAAGGATTTCTGAATATTTGCAAACTCTTGATGGAAGAGGAGCA 720
Db 661 ACCTAGCATCTGCAAAAGGATTTCTGAATATTTGCAAACTCTTGATGGAAGAGGAGCA 720
QY 721 AAGCAGATGGAATGCTCAAGATAATGAAGACATGTCCTCCACTCCATTTCTGTCGAT 780
Db 721 AAGCAGATGGAATGCTCAAGATAATGAAGACATGTCCTCCACTCCATTTCTGTCGAT 780
QY 781 TTGACACCATGATATAGTTAAGTATCTGCTCAAGAGTATTTGGAAGTTCAACCTCATG 840
Db 781 TTGACACCATGATATAGTTAAGTATCTGCTCAAGAGTATTTGGAAGTTCAACCTCATG 840
QY 841 TTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAATGGCAATTTGAAG 900
Db 841 TTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAATGGCAATTTGAAG 900

QY	901	TTGCCAAGAAATCATCAAAATATACAGGAACAGAAAGTCTGACTAAGGAAAAATCTTCA	960	QY	1981	AAGCAGATGTCCTCAGCTATGCTCTGTCTGTGGGAAATTCCTCACTGCGGAAATTCCTAT	2040
Db	901	TTGCCAAGGAAATCATCCAAATATACAGGAACAGAAAGTCTGACTAAGGAAAAATCTTCA	960	Db	1981	AAGCAGATGTCCTCAGCTATGCTCTGTCTGTGGGAAATTCCTCACTGCGGAAATTCCTAT	2040
QY	961	GTGAAACAGCTTTTCATAGTGTGTGTACCTAATGGCAAGAGCATTTGACCTAGTCAAAATTC	1020	QY	2041	TCGCTCATCTCAAGCCAGCGCTGCGGACAGACATGCGCTTACCAACCATCAGACCTC	2100
Db	961	GTGAAACAGCTTTTCATAGTGTGTGTACCTAATGGCAAGAGCATTTGACCTAGTCAAAATTC	1020	Db	2041	TCGCTCATCTCAAGCCAGCGCTGCGGACAGACATGCGCTTACCAACCATCAGACCTC	2100
QY	1021	TTCTTGATCAGAAATGTCATAAATCAACCAAGGAAGGATGGGCCACACTGGATTTAC	1080	QY	2101	CCATTGGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATACGAGGGTGGACGAT	2160
Db	1021	TTCTTGATCAGAAATGTCATAAATCAACCAAGGAAGGATGGGCCACACTGGATTTAC	1080	Db	2101	CCATTGGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATACGAGGGTGGACGAT	2160
QY	1081	ACTCTGCTGTGTACCAAGGTCACATTCGCTTGGTTCAGTTCTTACTGGATAATGGAGCTG	1140	QY	2161	GTCTCTGAAGGAAGACCGGAATTTTCTGAAGTTCTCATGAAGTTAGAAAGTGTCTCTGCA	2220
Db	1081	ACTCTGCTGTGTACCAAGGTCACATTCGCTTGGTTCAGTTCTTACTGGATAATGGAGCTG	1140	Db	2161	GTCTCTGAAGGAAGACCGGAATTTTCTGAAGTTCTCATGAAGTTAGAAAGTGTCTCTGCA	2220
QY	1141	ATATGAATCTAGTGTGTGATCCAGAGGTCCTAGTGTGAAAAGATGACAGACAT	1200	QY	2221	ACATTGAGCTGATGTCTCTGCAATCAAGTAACAGCAGTGGGTCTCTCTCACCTTCTTCT	2280
Db	1141	ATATGAATCTAGTGTGTGATCCAGAGGTCCTAGTGTGAAAAGATGACAGACAT	1200	Db	2221	ACATTGAGCTGATGTCTCTGCAATCAAGTAACAGCAGTGGGTCTCTCTCACCTTCTTCT	2280
QY	1201	GTTCGATGTGGCTTATGAAAAAGGCAATGATGCACTCTCTCTGAAAGCATTTATA	1260	QY	2281	CTTCTGATTCGCTGGTGAAACCGGGAGGACCTGGCCGGAGTCATGTGGCAGCATTAAGNA	2340
Db	1201	GTTCGATGTGGCTTATGAAAAAGGCAATGATGCACTCTCTCTGAAAGCATTTATA	1260	Db	2281	CTTCTGATTCGCTGGTGAAACCGGGAGGACCTGGCCGGAGTCATGTGGCAGCATTAAGNA	2340
QY	1261	AGAGACCAAGATGAATGGCTGTAAATGAATATCTCAGCCTGGAGGAGATGGCTCCT	1320	QY	2341	GTCTCTTTCGAATTCGAATATGCTCTAAATGCAAGGTCCTATGCTGCTTGTGCCAAAGTG	2400
Db	1261	AGAGACCAAGATGAATGGCTGTAAATGAATATCTCAGCCTGGAGGAGATGGCTCCT	1320	Db	2341	GTCTCTTTCGAATTCGAATATGCTCTAAATGCAAGGTCCTATGCTGCTTGTGCCAAAGTG	2400
QY	1321	ATGTGCTCTTCCATCAACCTTGGGGAAGATTAAGCATGACAAAAAGAGAGCAGATA	1380	QY	2401	CTGGAACAATATTCCTCTCAAGGTCGTCTTTTGGAGGAGATGAAGAGATCTTCAATACA	2460
Db	1321	ATGTGCTCTTCCATCAACCTTGGGGAAGATTAAGCATGACAAAAAGAGAGCAGATA	1380	Db	2401	CTGGAACAATATTCCTCTCAAGGTCGTCTTTTGGAGGAGATGAAGAGATCTTCAATACA	2460
QY	1381	TTCTCCTCTAAGAGCTGGATTCCTTCAATTCATTCCTCAGCTCTCAGAAATGAGT	1440	QY	2461	CACCCATTTGACAAATATGGCTATGTAATCCGATCCCATGAGCTCAATGCAATTTTCAATCT	2520
Db	1381	TTCTCCTCTAAGAGCTGGATTCCTTCAATTCATTCCTCAGCTCTCAGAAATGAGT	1440	Db	2461	CACCCATTTGACAAATATGGCTATGTAATCCGATCCCATGAGCTCAATGCAATTTTCAATCT	2520
QY	1441	TGCATGAGATTTGGCTCAGGTCCTTTTGGGAAGATATAAAGGACGATGACAGAAATA	1500	QY	2521	GCCAAATAGTAGCAGCTTTTGGAGCAGCAGCTGACAGCATTCGGGCTATACCTAAGGAG	2580
Db	1441	TGCATGAGATTTGGCTCAGGTCCTTTTGGGAAGATATAAAGGACGATGACAGAAATA	1500	Db	2521	GCCAAATAGTAGCAGCTTTTGGAGCAGCAGCTGACAGCATTCGGGCTATACCTAAGGAG	2580
QY	1501	AAATAGTGTCTATAAAGCTTATCGAGCCATACCTACTGCTCCAAAGTCAGATGTGATA	1560	QY	2581	AGTTTTCCTCCGAACTGACAGCAACGATTCGAACCAAGCTGCGCTTCCAACTATA	2640
Db	1501	AAATAGTGTCTATAAAGCTTATCGAGCCATACCTACTGCTCCAAAGTCAGATGTGATA	1560	Db	2581	AGTTTTCCTCCGAACTGACAGCAACGATTCGAACCAAGCTGCGCTTCCAACTATA	2640
QY	1561	TGTTTTCGAGAGTGTCCATCTCTGCGAGCTCAATCATCCCTGGTAAATCAGTTTG	1620	QY	2641	ACATTTTACTCTCAAGGTCCTCTTAAATTTGGGCTTGTCTTACTTGTCTTAAATTC	2700
Db	1561	TGTTTTCGAGAGTGTCCATCTCTGCGAGCTCAATCATCCCTGGTAAATCAGTTTG	1620	Db	2641	ACATTTTACTCTCAAGGTCCTCTTAAATTTGGGCTTGTCTTACTTGTCTTAAATTC	2700
QY	1621	TGGGTGCTTGTGAATGATCCAGCCAGTTTGGCAATGTGCACTCAATACATATCAGGGG	1680	QY	2701	CCCACTATTAGCAGGCTTGGATTTGCTGCTAAGGATATATATGCAAAAGAACCAAGACA	2760
Db	1621	TGGGTGCTTGTGAATGATCCAGCCAGTTTGGCAATGTGCACTCAATACATATCAGGGG	1680	Db	2701	CCCACTATTAGCAGGCTTGGATTTGCTGCTAAGGATATATATGCAAAAGAACCAAGACA	2760
QY	1681	GTTCTGTTTCCCTTCTATGAGAGAGAGAGATTTCTTGATTTGGAGCTTAAATTA	1740	QY	2761	GAAATGATATGAAGAAATGTTTAAATTTTAAATTTAAATTTAAATTTAAATTTAGATCGTACT	2820
Db	1681	GTTCTGTTTCCCTTCTATGAGAGAGAGAGATTTCTTGATTTGGAGCTTAAATTA	1740	Db	2761	GAAATGATATGAAGAAATGTTTAAATTTTAAATTTAAATTTAAATTTAGATCGTACT	2820
QY	1741	TTATTCAGTGTAGATGTGCAAGGCAATGAGTGTCTTCAACACTGACAGCAATTA	1800	QY	2821	TGGAAATGGAGCCCTAAGTCTGTGTGGACAGATAATTAATTTATGTTTCTCTGGGCTGAAAT	2880
Db	1741	TTATTCAGTGTAGATGTGCAAGGCAATGAGTGTCTTCAACACTGACAGCAATTA	1800	Db	2821	TGGAAATGGAGCCCTAAGTCTGTGTGGACAGATAATTAATTTATGTTTCTCTGGGCTGAAAT	2880
QY	1801	TACATCGTGTGAAACAGTCAATATTTCTTCTATGAGGATGGGATGCTGTGTGG	1860	QY	2881	ATGTAGACTTGTGTTTTCAGCAGTATGGGTTTATTTCTTAGAACATTTGTTCTTTT	2940
Db	1801	TACATCGTGTGAAACAGTCAATATTTCTTCTATGAGGATGGGATGCTGTGTGG	1860	Db	2881	ATGTAGACTTGTGTTTTCAGCAGTATGGGTTTATTTCTTAGAACATTTGTTCTTTT	2940
QY	1861	CAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAAGACAAATGACAAACAAC	1920	QY	2941	CTCATATTGTTACTCTTAGTGTTCACCTCTGTGATTAAGATTTCTTTGGTGAATAGAA	3000
Db	1861	CAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAAGACAAATGACAAACAAC	1920	Db	2941	CTCATATTGTTACTCTTAGTGTTCACCTCTGTGATTAAGATTTCTTTGGTGAATAGAA	3000
QY	1921	CTGGGAACCTCGTTCGATGGCTCTCTGAGGTTTCAACGAGTGCATCGGTACACCATCA	1980	QY	3001	AAAAAAGGAGGCGCGCGC 3025	
Db	1921	CTGGGAACCTCGTTCGATGGCTCTCTGAGGTTTCAACGAGTGCATCGGTACACCATCA	1980	Db	3001	AAAAAAGGAGGCGCGCGC 3025	

RESULT 3

US-09-947-199-3
; Sequence 3, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL DARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199-3

Query Match 82.8%; Score 2505; DB 9; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 ATGGGAAATTATAAATCTAGACCAACCCAACTTGTACTGATGAATGGAGAAAAAGTC 107
Db 1 ATGGGAAATTATAAATCTAGACCAACCCAACTTGTACTGATGAATGGAGAAAAAGTC 60

Qy 108 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGAAAAA 167
Db 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGAAAAA 120

Qy 168 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 227
Db 121 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180

Qy 228 AATTACGGCACTGAAATGGGCTGTCTACTTCAATTTATGTTGCAATTTGTGGAGGCAAG 287
Db 181 AATTACGGCACTGAAATGGGCTGTCTACTTCAATTTATGTTGCAATTTGTGGAGGCAAG 240

Qy 288 AATTCACATATTCGAATCTTTATGTTGAAGGGCTCGGCCATCTCGACTGCAAGAAAT 347
Db 241 AATTCACATATTCGAATCTTTATGTTGAAGGGCTCGGCCATCTCGACTGCAAGAAAT 300

Qy 348 GGATTTACAGCTTTGCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG 407
Db 301 GGATTTACAGCTTTGCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG 360

Qy 408 CTTACAGTGGAGCTGATATACAGCAGTTGGATAGGGTGGCTCACTGCCCTCCATATT 467
Db 361 CTTACAGTGGAGCTGATATACAGCAGTTGGATAGGGTGGCTCACTGCCCTCCATATT 420

Qy 468 GCTACAATAGTGGCCACCTAGAGGCTGTGATGCTGTGTCATGCAACATGGAGCTAATGTC 527
Db 421 GCTACAATAGTGGCCACCTAGAGGCTGTGATGCTGTGTCATGCAACATGGAGCTAATGTC 480

Qy 528 AATATCAAGATGAGTTTCTTTTCACTCCATTCGATGATGATGATGATGATGATGATGATG 587
Db 481 AATATCAAGATGAGTTTCTTTTCACTCCATTCGATGATGATGATGATGATGATGATGATG 540

Qy 588 CAGTAACTCGCTTCTTTTGAATTTGGTGTGATGATGATGATGATGATGATGATGATGATG 647
Db 541 CAGTAACTCGCTTCTTTTGAATTTGGTGTGATGATGATGATGATGATGATGATGATGATG 600

Qy 648 GATAGACCCCTCCACCTPAGCATCTGCAAAAGGATTTCTGAATATTGCAAAACTCTTGAT 707

Db 501 GATAGACCCCTCCACCTPAGCATCTGCAAAAGGATTTCTTGAATATTGCAAAACTCTTGATG 660
Qy 708 GAAGAAGGAGCAAAAGCAGATGTGAATGCTCAGATAATGAAGACCATGTCCACATCCAT 767
Db 661 GAAGAAGGAGCAAAAGCAGATGTGAATGCTCAGATAATGAAGACCATGTCCACATCCAT 720
Qy 768 TTCTGTTCTCGATTTGGACACCATGATATAGTTAACTATCTGTCGCAAGTCAATTTGGA 827
Db 721 TTCTGTTCTCGATTTGGACACCATGATATAGTTAACTATCTGTCGCAAGTCAATTTGGA 780
Qy 828 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAA 887
Db 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAA 840
Qy 888 GGCAGATTTGAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947
Db 841 GGCAGATTTGAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
Qy 948 GAAACATCTTTCAGTGAACAGCTTTTTCATAGTGTGTTACCTATGGCAAGAGCATTGAC 1007
Db 901 GAAACATCTTTCAGTGAACAGCTTTTTCATAGTGTGTTACCTATGGCAAGAGCATTGAC 960
Qy 1008 CTAGTCAAAATTTCTTGTATCAGATGTCATAAATCAACATCAACCAAGGAAGGATGG 1067
Db 961 CTAGTCAAAATTTCTTGTATCAGAAATGTATATAAATCAACATCAACCAAGGAAGGATGG 1020
Qy 1068 CACACTGGATTACACTCTGCTTGTACACAGGTCAATTCGCTGTGTTCACTTCTTACTG 1127
Db 1021 CACACTGGATTACACTCTGCTTGTACACAGGTCAATTCGCTGTGTTCACTTCTTACTG 1080
Qy 1128 GATAATGGAGCTGATATGAATCTAGTGGCTTGTATGCCAGAGGTCTAGTGGTGA 1187
Db 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTATGCCAGAGGTCTAGTGGTGA 1140
Qy 1188 GATGAGCAGACATGTTGATGTTGGCTTATGAAAAGGCGATGATGCCATTTGTACACTC 1247
Db 1141 GATGAGCAGACATGTTGATGTTGGCTTATGAAAAGGCGATGATGCCATTTGTACACTC 1200
Qy 1248 CTGAAGCATTTATAAGAGACCAAGATGAATTTGCCCTGTAATGAATATTTCTCAGCCTGA 1307
Db 1201 CTGAAGCATTTATAAGAGACCAAGATGAATTTGCCCTGTAATGAATATTTCTCAGCCTGA 1260
Qy 1308 GAGATGGCTCTTATGTTCTGTTCCATCACCTTGGGGAAGATTTAAAGCATGACAAA 1367
Db 1261 GAGATGGCTCTTATGTTCTGTTCCATCACCTTGGGGAAGATTTAAAGCATGACAAA 1320
Qy 1368 GAGAAGGCAGATATTTCTCCTCTAAGAGCTGGATTGCCCTTCACTTCCATCTTCAGCTC 1427
Db 1321 GAGAAGGCAGATATTTCTCCTCTAAGAGCTGGATTGCCCTTCACTTCCATCTTCAGCTC 1380
Qy 1428 TCAGAAATTGAGTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAAG 1487
Db 1381 TCAGAAATTGAGTTCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAAG 1440
Qy 1488 CGATGAGAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTTACCTGCTCCAG 1547
Db 1441 CGATGAGAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTTACCTGCTCCAG 1500
Qy 1548 TCAGATGTGATATGTTTTCGCCAGAGGTGTCCATTTCTCTGCCAGCTCAATCATCCCTGC 1607
Db 1501 TCAGATGTGATATGTTTTCGCCAGAGGTGTCCATTTCTCTGCCAGCTCAATCATCCCTGC 1560
Qy 1608 GTAAATTCAGTTTGTGGTGTCTTGAATGATCCAGCAGTTTGGCAATGTCACAT 1667
Db 1561 GTAAATTCAGTTTGTGGTGTCTTGAATGATCCAGCAGTTTGGCAATGTCACAT 1620
Qy 1668 TACATATCAGGGGGTCTCTGTTCTCCCTTCTATGAGCAAGAGGATTTCTTGAATTTG 1727
Db 1621 TACATATCAGGGGGTCTCTGTTCTCCCTTCTATGAGCAAGAGGATTTCTTGAATTTG 1680
Qy 1728 CAGTCTAAATTAATATTCAGTATGTTGCAAGGATGAGTACCTTCACAACTG 1787

Db 1681 CAGTCTAAATTAATTAATTCAGTAGATGTTGCCAAGGCGATGGAGTACCTTCAACACCTG 1740
QY 1788 ACACAGCCAAATTAATCATCGTGAATGAAAGTGCACAAATATCTCTCTATGAGGATGG 1847
Db 1741 ACACAGCCAAATTAATCATCGTGAATGAAAGTGCACAAATATCTCTCTATGAGGATGG 1800
QY 1848 CATGCTGTGGTGGCAGATTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1907
Db 1801 CATGCTGTGGTGGCAGATTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
QY 1908 ATGACAAAACAACTGGGAACCTCGTTGGATGCTCTGAGGTGTTTCAACGAGTGCAC 1967
Db 1861 ATGACAAAACAACTGGGAACCTCGTTGGATGCTCTGAGGTGTTTCAACGAGTGCAC 1920
QY 1968 CGGTACACCAATCAAGCAGATGTTTCAAGCTATGCTCTGCTGTCTGTGGGAAATTTCTCACT 2027
Db 1921 CGGTACACCAATCAAGCAGATGTTTCAAGCTATGCTCTGCTGTCTGTGGGAAATTTCTCACT 1980
QY 2028 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGCTGCGGAGCAGACATGGCTTACCAC 2087
Db 1981 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGCTGCGGAGCAGACATGGCTTACCAC 2040
QY 2088 CACATCAGACCTCCCATTTGGCTATTCATTCGAGCCCATATCATCTCTGCTGATACGA 2147
Db 2041 CACATCAGACCTCCCATTTGGCTATTCATTCGAGCCCATATCATCTCTGCTGATACGA 2100
QY 2148 GGGTGGAAACGATGCTCTGAAGGAAGACCCGAATTTTCTGAAGTTGTCATGAAGTTAGAA 2207
Db 2101 GGGTGGAAACGATGCTCTGAAGGAAGACCCGAATTTTCTGAAGTTGTCATGAAGTTAGAA 2160
QY 2208 GAGTGTCTCTGCAACATTTGAGTGTCTCTGATCAAGTAAACAGCAGTGGTCTCTC 2267
Db 2161 GAGTGTCTCTGCAACATTTGAGTGTCTCTGATCAAGTAAACAGCAGTGGTCTCTC 2220
QY 2268 TCACCT 2327
Db 2221 TCACCT 2280
QY 2328 GCAGCATTAAGAAGTGGTTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTT 2387
Db 2281 GCAGCATTAAGAAGTGGTTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTT 2340
QY 2388 TTGTCCCAAGTGGTGGCAATATTCCTCTCAAGTGTCTCTTGGAGGAGATGAAAGA 2447
Db 2341 TTGTCCCAAGTGGTGGCAATATTCCTCTCAAGTGTCTCTTGGAGGAGATGAAAGA 2400
QY 2448 AGTCTTCATACACACCCATTGACAAATATGCTATGCTATGCTATGCTATGCTATGCTATGCT 2507
Db 2401 AGTCTTCATACACACCCATTGACAAATATGCTATGCTATGCTATGCTATGCTATGCTATGCT 2460
QY 2508 CATTTTCATCTTGGGAAATAGTAGAGCTTTGAGGACGACG 2552
Db 2461 CATTTTCATCTTGGGAAATAGTAGAGCTTTGAGGACGACG 2505

RESULT 4

US-10-626-173-3
; Sequence 3, Application US/10626173
; Publication No. US20040110232A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/10/626,173
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: US/09/947,199A
; PRIOR APPLICATION NUMBER: US 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-10-626-173-3

Query Match 82.8%; Score 2505; DB 17; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 ATGGAATTAATAATCTAGACCAACCCAAACCTTGTACTGATGAATGGAGAAAGAAAGTC 107
Db 1 ATGGAATTAATAATCTAGACCAACCCAAACCTTGTACTGATGAATGGAGAAAGAAAGTC 60
QY 108 AGTGAATCATATGTTATCACAATAGAAAGATTAAGAAGATGACCTGCAGATCAAGGAAAAA 167
Db 61 AGTGAATCATATGTTATCACAATAGAAAGATTAAGAAGATGACCTGCAGATCAAGGAAAAA 120
QY 168 GAATCTGACAGAACTTAAGGAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA 227
Db 121 GAATCTGACAGAACTTAAGGAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA 180
QY 228 AATTACCGCACTGAAATGGGCTGTCTACTATTTCAATTTATGTTGCAATTTGTGGAGCAAG 287
Db 181 AATTACCGCACTGAAATGGGCTGTCTACTATTTCAATTTATGTTGCAATTTGTGGAGCAAG 240
QY 288 AAATCACATATTGCAACTCTTATGTTGAAGGGCTCGGCCATCTCGACTGACAGAAAT 347
Db 241 AAATCACATATTGCAACTCTTATGTTGAAGGGCTCGGCCATCTCGACTGACAGAAAT 300
QY 348 GGATTTACAGCCTTTGCAATTTAGCAGATTTACAAGGATAATGACAGAAATGATCACTTCTCTG 407
Db 301 GGATTTACAGCCTTTGCAATTTAGCAGATTTACAAGGATAATGACAGAAATGATCACTTCTCTG 360
QY 408 CTTACAGTGGAGCTGATATACAGCAGGTGGATACGGTGGCTCACTGCCCTCCATATT 467
Db 361 CTTACAGTGGAGCTGATATACAGCAGGTGGATACGGTGGCTCACTGCCCTCCATATT 420
QY 468 GCTCAATAGCTGCCACCTAGAGGCTGCTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
Db 421 GCTCAATAGCTGCCACCTAGAGGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 528 AATATTCAAGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 587
Db 481 AATATTCAAGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 588 CAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 647
Db 541 CAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 648 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAAATTTGCAAACTCTTGTGATG 707
Db 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAAATTTGCAAACTCTTGTGATG 660
QY 708 GAAGAGGCGACCAAGCAGATGTAATGCTCAAGATTAAGAAGCAGATGCTCCACCTCCAT 767
Db 661 GAAGAGGCGACCAAGCAGATGTAATGCTCAAGATTAAGAAGCAGATGCTCCACCTCCAT 720
QY 768 TTCTGTTCTCGATTTGGACACCATGATATAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 827
Db 721 TTCTGTTCTCGATTTGGACACCATGATATAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 828 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACACTGCTGCATGCTACAT 887
Db 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACACTGCTGCATGCTACAT 840
QY 888 GGCAGAAATTTGAAGTTGCAAGGAATCATCAATATFACAGAAACAGAAAGTCTGACTAAG 947

Db 841 GGCAAAATTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
Qy 948 GAAACAATCTTTCAGTGAACAGACTTTTCATAGTGTCTGTACCTATGCGCAAGAGCATTTGAC 1007
Db 901 GAAACAATCTTTCAGTGAACAGACTTTTCATAGTGTCTGTACCTATGCGCAAGAGCATTTGAC 960
Qy 1008 CTAGTCAAAATTTCTTGTGATCAGAAATGTCTATAAAATCAACACCAAGAGGATGGG 1067
Db 961 CTAGTCAAAATTTCTTGTGATCAGAAATGTCTATAAAATCAACACCAAGAGGATGGG 1020
Qy 1068 CACACTGGATTACACTCTGCTGTGTACACGGTTCACATTCGCTGTGTTCAGTTCTTACTG 1127
Db 1021 CACACTGGATTACACTCTGCTGTGTACACGGTTCACATTCGCTGTGTTCAGTTCTTACTG 1080
Qy 1128 GATTAATGGAGCTGATATGAATCTAGTGGCTGTGATCCAGCAGGTCTAGTGGTGAAGAA 1187
Db 1081 GATTAATGGAGCTGATATGAATCTAGTGGCTGTGATCCAGCAGGTCTAGTGGTGAAGAA 1140
Qy 1188 GATCAGCAGACATCTTTGATGTGGCTTATGAAAAGGGCATGATGCCATTTGTCACTC 1247
Db 1141 GATCAGCAGACATCTTTGATGTGGCTTATGAAAAGGGCATGATGCCATTTGTCACTC 1200
Qy 1248 CTGAAGCATTTAAGAGACACCAAGATGAATTCCTGTGAATGAATATTTCTCAGCCTGGA 1307
Db 1201 CTGAAGCATTTAAGAGACACCAAGATGAATTCCTGTGAATGAATATTTCTCAGCCTGGA 1260
Qy 1308 GGAGATGGCTCCTATGTCTGTTCATCACCTTGGGGAGATTTAAAAGCATGACAAA 1367
Db 1361 GGAGATGGCTCCTATGTCTGTTCATCACCTTGGGGAGATTTAAAAGCATGACAAA 1320
Qy 1368 GAGAAGCAGATATTTCTCCTTAAGAGCTGGATTCCTTCATATTCATCTTCAGCTC 1427
Db 1321 GAGAAGCAGATATTTCTCCTTAAGAGCTGGATTCCTTCATATTCATCTTCAGCTC 1380
Qy 1428 TCAGAAATTCAGTTCATAGATATTTGGCTCAGGTTCTTTGGGAAAGTATATAAGGA 1487
Db 1381 TCAGAAATTCAGTTCATAGATATTTGGCTCAGGTTCTTTGGGAAAGTATATAAGGA 1440
Qy 1488 CGATGCAGAAATAAATAGTGGCTATAAAAGCTTATCGAGCCAAATACCTACTGCTCCAG 1547
Db 1441 CGATGCAGAAATAAATAGTGGCTATAAAAGCTTATCGAGCCAAATACCTACTGCTCCAG 1500
Qy 1548 TCAGATGTGATATGTTTTGCCAGAGGTCTCAATTCCTGCGAGTCAATCATCCCTGC 1607
Db 1501 TCAGATGTGATATGTTTTGCCAGAGGTCTCAATTCCTGCGAGTCAATCATCCCTGC 1560
Qy 1608 GTAAATTCAGTTTGGGTGCTTCTGAAATGATCCAGCCAGCTTGGCAATTCCTACTCA 1667
Db 1561 GTAAATTCAGTTTGGGTGCTTCTGAAATGATCCAGCCAGCTTGGCAATTCCTACTCA 1620
Qy 1668 TACATATCAGGGGTCTCTGTTCTCCCTCTCATGAGCAGAGAGGATTCCTGATTTG 1727
Db 1621 TACATATCAGGGGTCTCTGTTCTCCCTCTCATGAGCAGAGAGGATTCCTGATTTG 1680
Qy 1728 CAGTCTAAATTAATTTGAGTAGATGTTGCCAAAGGATGAGTACCTTCACAACTG 1787
Db 1681 CAGTCTAAATTAATTTGAGTAGATGTTGCCAAAGGATGAGTACCTTCACAACTG 1740
Qy 1788 ACACAGCAATTTACATCTGACTTGAACAGTCAATATTTCTCTATCAGAGATGG 1847
Db 1741 ACACAGCAATTTACATCTGACTTGAACAGTCAATATTTCTCTATCAGAGATGG 1800
Qy 1848 CATGCTGTGTGGCAGATTTGGAGATCAAGATTTCTACGTCTCTGGATGAAGCAAC 1907
Db 1801 CATGCTGTGTGGCAGATTTGGAGATCAAGATTTCTACGTCTCTGGATGAAGCAAC 1860
Qy 1908 ATGACAAAACCTGGGACCTCGTTGGATGGCTTGGAGTGTTCACGAGTGCAC 1967
Db 1861 ATGACAAAACCTGGGACCTCGTTGGATGGCTTGGAGTGTTCACGAGTGCAC 1920
Qy 1968 CGGTACACCATCAAGCAGATGCTTCAGCTATGCTGTGTCTGTGGGAAATTTCTCACT 2027
Db 1921 CGGTACACCATCAAGCAGATGCTTCAGCTATGCTGTGTCTGTGGGAAATTTCTCACT 1980

Qy 2028 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGGAGCAGACATGGCTTACCAC 2087
Db 1981 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGGAGCAGACATGGCTTACCAC 2040
Qy 2088 CACATCAGACTCCCATTTGCTATTCATTCCTCAAGCCCATATCATCTCTGCTGATACA 2147
Db 2041 CACATCAGACTCCCATTTGCTATTCATTCCTCAAGCCCATATCATCTCTGCTGATACA 2100
Qy 2148 GCGTGGAAACGATGCTCTGAAAGAACCCGAAATTTCTGAAAGTTCTCATGAAGTTAGAA 2207
Db 2101 GCGTGGAAACGATGCTCTGAAAGAACCCGAAATTTCTGAAAGTTCTCATGAAGTTAGAA 2160
Qy 2208 GAGTGTCTCTGCAACATTCAGTGTCTCTGCAATCAAGTAAACAGAGTGGTCTCTC 2267
Db 2161 GAGTGTCTCTGCAACATTCAGTGTCTCTGCAATCAAGTAAACAGAGTGGTCTCTC 2220
Qy 2268 TCACCTTCTTCTTCTTCTGATTCCTGTAACCGGGAGGACCTGCGCGAGTCAATG 2327
Db 2221 TCACCTTCTTCTTCTTCTGATTCCTGTAACCGGGAGGACCTGCGCGAGTCAATG 2280
Qy 2328 GCAGCATTAAGAAGTCTGTTTTCGAATTCGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2387
Db 2281 GCAGCATTAAGAAGTCTGTTTTCGAATTCGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2340
Qy 2388 TTGTCCTCAAGTCTGCAACATTCCTCTCAAGTCTGCTTTGGAGGAGATGAAGA 2447
Db 2341 TTGTCCTCAAGTCTGCAACATTCCTCTCAAGTCTGCTTTGGAGGAGATGAAGA 2400
Qy 2448 AGTCTTCAATACACACCCATTGACAAATATGCTATGCTATGCTATGCTATGCTATGCTATG 2507
Db 2401 AGTCTTCAATACACACCCATTGACAAATATGCTATGCTATGCTATGCTATGCTATGCTATG 2460
Qy 2508 CATTTTCATTTCTTGGCGAATATAGTAGCAGCTTTGAGGACAGAGC 2552
Db 2461 CATTTTCATTTCTTGGCGAATATAGTAGCAGCTTTGAGGACAGAGC 2505

RESULT 5
US-09-947-199-7
; Sequence 7, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3026
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(2565)
US-09-947-199-7

Query Match 66.4%; Score 2010; DB 9; Length 3026;
Best Local Similarity .80.8%; Pred. No. 0;
Matches 2425; Conservative 0; Mismatches 570; Indels 13; Gaps 6;
Qy 16 CGGCCCTTGGAGAAAGGAGAACTTATAATAATGGAAATTTATAATCTAGACCAACC 75
Db 29 CAGCAGGAGAGAAAGCAAGACTTCTTTAAATGGGAAATTACAAATCCAGACCAAC 88

QY	76	AAACTTGTACTGATGAATCGAAGAAAAAGTCACTGAATCATATGTTATCAATAGAAA	135
DB	89	AGACTTGTCTGATGAATGAAGAGAAAGTTAGTGAATCTTCGCTATTATCATAGAAA	148
QY	136	GATTAGAAAGATGACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAGGAATATATTG	195
DB	149	GGCTGGAGGATAA CCTGCAGATCAAAAGAAAAATGAATTTCAAGNACTAAGGCACATCTTTG	208
QY	196	GCTCTGATGAAGCCTTCAGTTAAAGTCAATTTAAATTAACCGCACTGAAAAATGGCGCTGCTC	255
DB	209	GCTCTGATGAAGCCTTCAGTTAAAGTCAAGTTAAATTAACCGCAGAGCGTGGCGCTGCTCC	268
QY	256	TACTTCAATTTATGTTGCATTTGTGGAGGCCAAAGAAATTCACATATTCGAACTCTTATGTTGA	315
DB	269	TGCTACACCTCTGCTGTGTCTGTGGCGGCAACAAGTCCATATCGTGCCCTTATGTTAA	328
QY	316	AAGGCGTCGGCCCATCTCGACTGACAGAAATGGATTTACAGCCTTGCAATTTAGCAGTTT	375
DB	329	AAGGCGTCGGCCCATCTCGACTGACAGAAATGGGTTTCCAGCTCTGCACCTGGCCGCTTT	388
QY	376	ACAAGGATAATGCGAAGATTCATCTCTCTGCTTCACAGTGGAGCTCATATACAGCAGG	435
DB	389	ACAAGGACAGCCCGAATTTATCACTTCACTGTTGCAAGCGGACGATGTTCAAGCAAG	448
QY	436	TTTGGATACGGTGGCCTCACTGCCCTCCATATTGTTCTAATAAGTCTGGCCACCTAGAGCGTG	495
DB	449	TGGGATACGGTGGCCTCAAGCCCTCCACATAGCTGCAATAGCTGGAACCCACAGAGCGTG	508
QY	496	CTGATGCTGTTGCAATGAGCGACTTAATGTCATTAATTCAGATGSCAGTTTTTTTCACTC	555
DB	509	CAGAAAGTGTCTCTCAACATGCGGGCCAATGTGAATGTTCAAGATCCGCTCTCTTCAACC	568
QY	556	CAATGCAATTTGCAAGCGTACTATGGACATGAACAGGTAACTCGCCTCTTTTGAATTTG	615
DB	569	CACCTGCATTTGCAAGCTACTATGGGCAAGCAGGTAAACCAAGTGTCTTTTGAAGTTG	628
QY	616	GTGCTGATGTAATATGTAAGTGTGAAGTTGGAGATAGACCCCTCCACCTACGATCTCGAA	675
DB	629	GTGCTGATGTCAATGTAAGCGGTGAAGTTGGGACAGGCGCTCTGCACCTGGCCCTCTCGAA	688
QY	676	AAGGATCTCTGAATATTGCAAAACTCTTGATGGAAGAGGACGAAAGACAGATGTGAATG	735
DB	689	AGGCGTTCTTTCAACATTTGTGNACTCTCTGTAGAAGAGGAGCAACACAGATGTGAACG	748
QY	736	CTCAAGATAATGAAGACATGTGCCACTCAATTTCTGTCTCGATTTGGACACCATGATA	795
DB	749	CTCAGGACAATGAAGACCAAGTCGCTCTGCACTTCTGTCTCGATTTGGACACCAACAATA	808
QY	796	TAGTTAAGTATCTGTCGCAAGTGATTTGGAAGTTCACCTCATGCTGTGTTAATCTATG	855
DB	809	TAGTAGCTACCTGCTCCAGATGACCTTAGAGTCCAGCTCCAGCTCATTTAACTATATG	868
QY	856	GAGATACCCCTTTACCTGTGCGATGCTACAATGGCAAAATTTGAAGTTCCTCAAGAAATCA	915
DB	869	GTGACACTCTTTTGACCTGGCATGCTACAAATGGAAATTTTGAAGTTCCTCAAGAAATG	928
QY	916	TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAAACATCTTCAGTGAACAGCTTTTC	975
DB	929	TCCAGGTAACAGGAACTGAAATGCTGACTAAGGAAAACATCTTCAGCAGACAGCTTTTC	988
QY	976	ATAGTCTTTGTACCTATGGCAAGAGACTTACCTAGTCAAAATTTCTTCTGATCAGAATG	1035
DB	989	ACAGTGTGTGTACTATGGCAAGAACATTTGACCTGGTCAAAATTTCTTCTGATCAGAATG	1048
QY	1036	TCATAAACAATCAACCAAGGAAGGATGGGCAACTCTGGATACACTCTGCTTGTCTACC	1095
DB	1049	CTGTGAACAATTAACCAACGAGGAAGAGATGGGCAACAGGAATTTGCACTCTCTGCTACC	1108
QY	1096	ACGCTCACATTCGGCTGGTTCAAGTCTTCTTACTTGGAATATGGAGCTGATATGAATCTAGTGG	1155
DB	1109	ACGGCCATATCCGCTTGGTTCAAGTCTCTACTCTTGATATGGTGCAGATATGATCTTCTCG	1168

QY	1156	CTTTGTGATCC	CAGAGTCTAGTGGTGA	AAAGATGAGCAGACATGTTTGTGATGTGGCGTT	1211
Db	1169	CTTTGTGATCC	CAGAGTCTAGTGGTGA	AAAGATGAGCAGACATGTTTGTGATGTGGCGTT	1228
QY	1216	ATGAAAAGGCG	ATGATGCCATTGTCACCTCGAAGCATTA	TAAGAGACACAAGATG	1275
Db	1229	ACGAAAAGGAC	ATGATGCCATTGTTACCTCTGAAGCATTA	CAAGAGACCCACAGAGG	1288
QY	1276	AAATGCCCTCT	TAATGAATATTTCTCAGCCTGGAGAGATGGCTCTATGTGTCTGTTCCAT	1335	
Db	1289	AGCTGCCATGT	AACGAATATTTCCAGCCTGGAGAGATGGCTCTATGTGTCTGTTCCAT	1348	
QY	1336	CACCTTCGGGA	AGATTAAGAAGCATGACAAAAGAGAGAGCAGATATTCCTCTCCTCAAGAG	1395	
Db	1349	CCCTTCGGGA	AGATTAAGAAGCATGACAAAAGAGAGAGCAGATATTCCTCTCCTCAAGAG	1408	
QY	1396	CTGGATTGCC	TTACATTTCCATCTTTCCAGTCTTCAGAAATTCAGTTTCCATGAGATTAATG	1455	
Db	1409	CTGAATACCT	CCCGCTTCCATCTCCAACTCTCCGAATTCGAGTTCCACGAGATTAATG	1468	
QY	1456	GCTCAGGTTCT	TTTGGAAAGTATATAAAGGACGATGACAGAAATAAAATAGTGCCTATAA	1515	
Db	1469	GCTCGGTTCT	TTTGGAAAGTCTATAAAGGCGGATGACAGAAATAAAATAGTGCCTATAA	1528	
QY	1516	ACGTTATCGAG	CCAAATACCTACTGCTCCAAAGTCAGATGTCGATATGTTTTCGCGAGAGG	1575	
Db	1529	AACGATCCGAG	CCAAACACTATCTGCTCCAAAGTCAGATGTCGATATGTTTTCGCGAGAGG	1588	
QY	1576	TGTCCATTCT	CTGGCAGTCAATCATCCCTCGCTAAATTCAGTTTGTGGGTGCTTCTGTGA	1635	
Db	1589	TGTCATTCT	CTGGCAGTCAACACCCCTCGCTGTTTCAGTTTGTGGGTGCTTCTGTGTG	1648	
QY	1636	ATGATCCAGC	CGAGTTTGCCATTGTGCATCAATACATATCAGGGGTTCTCTGTTCTCC	1695	
Db	1649	ATGACCCAGT	CGAGTTTGCCATTGTGCATCAATACATATCAGGGGTTCTCTGTTCTCC	1708	
QY	1696	TCCTTCATGAG	CAGAAGAGATTTCTGATTGTCAGTCTAAATTAATATTCAGTATAGT	1755	
Db	1709	TGCTTCATGAG	CAGAAGAGATTTCTTGACTTGCGATCTAAATTAATATTCAGTATAGT	1768	
QY	1756	TTGCCAAAGGA	ATGAGTACCTTCACAACTGACACAGCCATTTATACATCGTGACTTGA	1815	
Db	1769	TTGCCAAAGGA	ATGAGTACCTTCACAACTGAGTCTGACCGCCATCATACACCGGACCTTGA	1828	
QY	1816	ACAGTCAAA	TATTTCTCTATGAGATGGGATGCTGTGTGTGGCAGATTTTGGAGAA	1875	
Db	1829	ACAGCCACAT	TATTTCTGCTCTATGAGATGGGATGCGCATGCTGTGTGTGGCAGATTTTGGAGAA	1888	
QY	1876	CAAGATTTCT	ACAGTCTCTGGATGAGACAAATGACAAACAACTGGGAACCTCCGTT	1935	
Db	1889	CAAGATTTCT	ACAGTCTCTGGATGAGACAAATGACAAACAACTGGGAACCTCCGCGCT	1948	
QY	1936	GGATGGCTCT	GTGAGTGTTCACGAGTGCATCGGTACACCATCAAAAGCAGATGCTTCA	1995	
Db	1949	GGATGGCTCT	GTGAGTGTTCACAGTGCAGAGATACACCATCAAAAGCAGATGCTTCA	2008	
QY	1996	GCTATGCTCT	GTGTGGGAAATTTCTCATCTGGCGAAATTCATTCGCTCATCTCAAGC	2055	
Db	2009	GTTACTCCTCT	GTGTGGGAGTCTCTCATCTGGAGAAATTCATTCGCTCATCTCAAGC	2068	
QY	2056	CAGCGGTCTG	CGCAGCAGATGAGCTTTACACACATCAGACCTCCCATTCGCTATTCCA	2115	
Db	2069	CAGCGGTCTG	CGCAGCAGATATGAGCTTTACACACATCAGACCGCCCATTCGCTATTCCA	2128	
QY	2116	TTCCCAAGCCC	ATATCATCTCTGCTGATACAGGGGTGGAAACGATGTCCTGAAGGAAGAC	2175	
Db	2129	TCCCAAGCCCA	TTCTATCTGATACCGGGGTGGAAATGATGCTCTCTGAAGGAAGAC	2188	
QY	2176	CGGAATTTCT	GAGTTGTGATGAAGTTAGAGAGTGTCTCTGCAACATTTAGAGCTCATGT	2235	
Db	2189	CAGAGTTCTCT	GAGTTGTGATGAAGTTAGAGAGTGTCTCTGCAACATTTAGAGCTCATGT	2248	
QY	2236	CTTCCTGCAT	CAAGTAACAGCAGTGGGTTCTCTCTCACCTTCTTCTTCTGATTCGCTGG	2295	

Db	869	GTGACACTCCTTTGACCTGGCATGCTACAAATGGAATTTTGAAGTTGCGCAAGAAATTTG	928	1949	GGATGGCCCTGAGGTGTTTACACAGTGCACGAGATACACCATCAAGGCTGATGCTTCA	2008
Qy	916	TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTCAAGTAAACAGCTTTTC	975	1996	GCTATGCTGTGTCTGTGGGAAATTTCTCACTGGGGAAATTTCCATTTGCTCATCTCAAGC	2055
Db	929	TCCAGGTAACAGAACTGAAAGTCTGACTAAGGAAACATCTTCAAGGAGACAGCTTTTC	988	2009	GTTACTCCCTGTGTGTGGGAGCTCTCACTGGGAAATTTCCATTTGCTCATCTCAAGC	2068
Qy	976	ATAGTCTGTGACCTATGGCAAGACATTTGACCTAGTCAAAATTTCTTCTTGATCAGAAATG	1035	2056	CAGGGCTGGGAGACAGACATGGCTTACCACATCAGACCTCCCATTTGGCTATTCCA	2115
Db	989	ACAGTGTGTGATGATGGCAAGAACATTTGACCTGGTCAAAATTTCTTCTTGATCAGAAATG	1048	2069	CAGCGCTGAGCAGCAGATATGGCGTATCAACACATCAGACCGGCTATTCGCTATTCCA	2128
Qy	1036	TCATTAACATCAACCAACGAGGAGGATGGGACATGATTTACATCTGCTTGTCTACC	1095	2116	TTCCAGCCCATATCATCTCTGCTGATACGAGGAGGAAACGATGCTCTGAAGAAAGAC	2175
Db	1049	CTGTGAACATTAACCAACGAGGAGGATGGGACATGATTTACATCTGCTTGTCTACC	1108	2129	TTCCAGCCCATATCATCTCTGCTGATACGAGGAGGAAACGATGCTCTGAAGAAAGAC	2188
Qy	1096	ACGGTCAATTCGGCTGCTTCACTTCTTCTGATATGAGGAGTATGATGATGATGATG	1155	2176	CCGAAATTTCTGAAGTTGTCATGAAGTTAGAGAGTCTCTGCAACATTTGAGCTGATGT	2235
Db	1109	ACGGCCATATCCGCTGCTTCACTTCTTCTGATATGAGGAGTATGATGATGATGATG	1168	2189	CAGAGTTCTCTGAAGTCGTTAGCAAACTGGAGAGTGCCCTATGCAATGTGAGCTCATGT	2248
Qy	1156	CTTGTGATCCGAGCAGTCTAGTGGTGAAGAGATGAGCAGATGTTTGTGAGTGGCTT	1215	2236	CTCTGCTCATCAAGTAAACAGCAGTGGGTCTCTCTCACCTTTCTTCTTCTGATGCTCTGG	2295
Db	1169	CTTGTGATCCGAGCAGTCTAGTGGTGAAGAGATGAGCAGATGTTTGTGAGTGGCTT	1228	2249	CTCCAGCATCAAGTAAACAGCAGTGGGTCTCTCTCACCTTTCTTCTTCTGATGCTCTGG	2308
Qy	1216	ATGAAAAAGGATGATGCCATTTCTCACTCTTGAAGCATTATTAAGAGACCAACAAGATG	1275	2296	TGAACCGGGAGGAGACCTGGCCGAGTCTATGTGGCAGCATTAAGAGTGGTTTCGAAATGG	2355
Db	1229	ACGAGAAAGGATGATGCCATTTCTCACTCTTGAAGCATTATTAAGAGACCAACAAGATG	1288	2309	TGAGCCGGGAGGAGCCTGGCCGAGCAGCTGGCAGCCTTACGAGCCTTTTGTGAGTTGG	2368
Qy	1276	AATGCCCTGTAATTAATTTCTCAGCCTGGAGAGATGGCTCTATGCTGCTGCTTCCAT	1335	2356	AATATGCTCTAAATGCAAGTCTCTATGCTGCTTTGTCCCAAGTGTGCAACATATTCTCT	2415
Db	1289	AGCTGCCATGTAACGAATTTCCAGCCTGGAGGAGTGGCTCTATGCTGCTGCTTCTT	1348	2369	AGTATGCCCTTAATGCAAGTCTCTATGCTGGGTGGTCCCAAGTGTGGAACACACTCTA	2428
Qy	1336	CACCTCTGGGAGATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATG	1395	2416	CTCAAGTCTGTCTTTGGAGGAGATGAAAGAGTCTTCAATACACACCCATTTGACAAAT	2475
Db	1349	CCCCCTTGGGAGATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATG	1408	2429	ATCCGGCCTGTCTTTGGAGGAGATGAATAGGAGCACCCAGTATTTCAACTGTTGCAAAAT	2488
Qy	1396	CTGATGCTTCAATTTCCATTTCCATTTCCATTTCCATTTCCATTTCCATTTCCATTTCC	1455	2476	ATGGCTATGATCCGATCCCATGAGTCTCAATGCTTTTCTTCTTCTGCGCAATATTAGCA	2535
Db	1409	CTGAATACCTCCGCTTCCATTTCCATTTCCATTTCCATTTCCATTTCCATTTCCATTTCC	1468	2489	ACGGCTATGCTGTGATFCCCATGAGCCTGACGACCTTCACTCCGCGCAAGACAGCA	2548
Qy	1456	GCTCAGGTTCTTTGGGAGATTAAGAGATGATGATGATGATGATGATGATGATGATG	1515	2536	GCTTTGAGGAGCAGCAGCTGACAGCTTCCGCGTATACCTTAAGAGAGTGTTCCTCCGAA	2595
Db	1469	GCTCGGTTCTTTGGGAGATTAAGAGATGATGATGATGATGATGATGATGATGATGATG	1528	2549	ACTTTGGAGCAGCACTGACAG-GTCTGGCATACACTTAAGGGGCTCTCCCATCAGG	2607
Qy	1516	AACGTTATCGAGCAGTCTGCTTCACTCTCAGAGTGTGATGATGATGATGATGATGATG	1575	2596	CTGACAGCAAGATTTCCACACCGCAAGTGGCTTCCAACTATACATTTTACTCTCAAA	2655
Db	1529	AACGATACCGAGCAGCAGTCTGCTTCACTCTCAGAGTGTGATGATGATGATGATGATG	1588	2608	CTGACAGCAGTATTTTACCCATGGCAGGCTTCTTCCAAATTAATACGCGCTGCGCTCTG	2667
Qy	1576	TGTCATTTCTGAGCAGTCTCAATCATCTCGCTGATTAATTCAGTTTGGGTGCTGCTTGA	1635	2656	AGGTCTCTTAAATTTGGGCTTGTGTTTACTTGTCTTATTTAAATTTCCCACTATTAGCAGG	2715
Db	1589	TGTCATTTCTGAGCAGTCTCAATCATCTCGCTGATTAATTCAGTTTGGGTGCTGCTTGA	1648	2668	AGGT-TTCTTCAATCGTCTGCTTATTTCTAAGCTGTTTAAATTTCCCTTCTACAGCAG	2726
Qy	1636	ATGATCCAGCAGTCTGCTTCACTCTCAATACATATCAGGGGTTCTCTGTTCTCTCC	1695	2716	CTTTGGATTTGCTGCTTAAGAAATTAATGCAAAAGAACCAAGACAGATGATATGAGA	2775
Db	1649	ATGATCCAGCAGTCTGCTTCACTCTCAATACATATCAGGGGTTCTCTGTTCTCTCC	1708	2727	GCTTTGATCTATGCC--AAGCCTGAAGTGTCAAAGAGCAGATACAGAAATGTCATGAGA	2784
Qy	1696	TCCTTTCATGAGCAGAGGATTTCTGATTTGAGTCTAAATTAATTTATTTGAGTATG	1755	2776	ATTGTTTTTAAATTTGTAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA	2835
Db	1709	TGCTTTCATGAGCAGAGGATTTCTGATTTGAGTCTAAATTAATTTATTTGAGTATG	1768	2785	ATTGTTTTTAAATTTGTAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA	2840
Qy	1756	TTGCCAAAGGATGAGTATGCTTCACTCAACCTGACACAGCCTTATATACATCGTACCTGA	1815	2836	AGTCTGTGGTGGACAGATAAATTAATTTGTTTCTGGGTGATTAATTTAGACTTTGTTT	2895
Db	1769	TTGCCAAAGGATGAGTATGCTTCACTCAACCTGACACAGCCTTATATACATCGTACCTGA	1828	2841	TGTGTAGATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	2900
Qy	1816	ACAGTCACAATATTTCTTCTATGAGGATGGGATGCTGTGGTGGCAGATTTTGGAGAT	1875	2896	TGACAG-CTATGGGTTTATTTCTTAGACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	2954
Db	1829	ACAGTCACAATATTTCTTCTATGAGGATGGGATGCTGTGGTGGCAGATTTTGGAGAT	1888	2901	TGAGGGTCTCCTGTGGGTTTCTTAAACCAAGCTGGCTGATTTATCTCTCTCTTCTCTCTT	2960
Qy	1876	CAAGATTTCTCAGTCTCTGGATGAGCAGATGATGATGATGATGATGATGATGATGATG	1935	2955	TCTAGTGTTCACCTCTGTGATTAAGATTTCTTTGGTGAATAGAAAAAAGAAAAAAGAAAA	3014
Db	1889	CAAGATTTCTCAGTCTCTGGATGAGCAGATGATGATGATGATGATGATGATGATGATG	1948	2961	---GTTTGTACTTCTGTGATTAAGTCTCTTGGTGTCTAGAAAAAAGAAAAAAGAAAA	3016
Qy	1936	GGATGGCTCCTCAGTGTTCACGAGTGCATCGGTACACCATCAAGAGAGATGTTCTTCA	1995	3015	AGGGGGG 3022	
				3017	GGGGGGG 3024	

RESULT 7

US-09-947-199-9
; Sequence 9, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: THEREFOR
; FILE REFERENCE: MN-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199-9

Query Match 62.6%; Score 1892.2; DB 9; Length 2505;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

QY	48	ATGGGAAATTAATAAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGATC	107
DB	1	ATGGGAAATTAATAAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAGATC	60
QY	108	AGTGAATCATATGTTATCAATAGAAAGATAGAAAGATGACCTCAGATCAAGGAAAAA	167
DB	61	AGTGAATCTTACGCTATTATCATAGAAAGATAGAAAGATGACCTCAGATCAAGGAAAAA	120
QY	168	GAATGACGAACTAAGAAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA	227
DB	121	GAATTTCAAGAACTAAGGACATCTTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA	180
QY	228	AATTACGACATGAATAGGCTGCTCTACTTCTATTTATGTTGATTTGCTGTTGGGGAAG	287
DB	181	AATTACGACATGAATAGGCTGCTCTACTTCTATTTATGTTGATTTGCTGTTGGGGAAG	240
QY	288	AAATCATATATTCGAATCTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAAT	347
DB	241	AAATCATATATTCGAATCTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAAT	300
QY	348	GGATTTACAGCTTGCATTTAGCAGTTTACAGGATATGCGATTTGATCACTTCTCTG	407
DB	301	GGATTTTCCAGCTTGCATTTAGCAGTTTACAGGATATGCGATTTGATCACTTCTCTG	360
QY	408	CTTTCAGATGGAGCTGATATACAGAGGTTGATACGGTGGCTCAGTCCCTCCATATT	467
DB	361	TTGCACAGGGACAGATGTTCAAGAGTGGATACGGTGGCTCAGTCCCTCCATATT	420
QY	468	GCTCAATAGTGGCCACTAGAGGCTGCTGATGCTGTTGCAACATGAGAGCTAATGTC	527
DB	421	GCTGCAATAGTGGCCACTAGAGGCTGCTGATGCTGTTGCAACATGAGAGCTAATGTC	480
QY	528	AATATTCAAGATGCACTTTTTCATCCATGCTATATGACGGTACTATGGACATGAA	587
DB	481	AATTTTCAAGATGCACTTTTTCATCCCTGCTATGACGGTACTATGGACATGAA	540
QY	588	CAGGTAACCTGCGCTCTTTTGAATTTGGTGTGCTGATGTAATGTAAGTGGTGAAGTTGA	647
DB	541	CAGGTAACCTGCTCTTTTGAATTTGGTGTGCTGATGTAATGTAAGTGGTGAAGTTGG	600

QY	648	GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAAACTCTTGATG	707
DB	601	GACAGGCCCTCTGCACCTGGCCCTCTGCAAAAGGCTTTCTTCAACATTTGTAAGACTCTGTA	660
QY	708	GAAGAGGCGAGCAAAAGCAGATGTAATGCTCAAGATATGAAGCAATGTCCTCCATCCAT	767
DB	661	GAAGAGGCGAGCAAAAGCAGATGTAATGCTCAAGCAATGAAGCAATGTCCTCCATCCAT	720
QY	768	TTCTGTTCTCGATTTTGGACACCATGATATAGTTAAAGTATCTGCTGCAAAAGTGAATTTG	827
DB	721	TTCTGTTCTCGATTTTGGACACCATGATATAGTTAAAGTATCTGCTGCAAAAGTGAATTTG	780
QY	828	GTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTACACCTGGCATGCTAGAT	887
DB	781	GTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTACACCTGGCATGCTAGAT	840
QY	888	GGCAAAATTTGAAAGTTGCCAAGGAAATCATCAAAATATCAGGAACAGAAAGTCTGACTAAG	947
DB	841	GGAAATTTTGAAGTTGCCAAGGAAATTTGCCAGGTAACAGGAATCTGAAAGTCTGACTAAG	900
QY	948	GAAACATCTTCAGTGAACACAGCTTTTTCATAGTGTGTTGTAATGGAAGCAAGATTTGAC	1007
DB	901	GAAACATCTTCAGTGAACACAGCTTTTTCATAGTGTGTTGTAATGGAAGCAAGATTTGAC	960
QY	1008	CTAGTCARATTTCTTCTTGATCAGAAATGTCATAACATCAACCAAGGAGGATGGG	1067
DB	961	CTGTCARATTTCTTCTTGATCAGAAATGTCATAACATCAACCAAGGAGGATGGG	1020
QY	1068	CACACTGGATTAACACTCTGCTTGTGCTTACCAAGGTCATTCGCTGGTTCAGTCTTACTG	1127
DB	1021	CACACTGGATTAACACTCTGCTTGTGCTTACCAAGGTCATTCGCTGGTTCAGTCTTACTG	1080
QY	1128	GATATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCAAGGCTGCTAGTGGTGAAGAA	1187
DB	1081	GATATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCAAGGCTGCTAGTGGTGAAGAA	1140
QY	1188	GATGACACACATGTTTTCATGTTGGCTTTATGAAAGGCGCATGATGCCATTTGTCACTC	1247
DB	1141	GATGACACACATGTTTTCATGTTGGCTTTATGAAAGGCGCATGATGCCATTTGTCACTC	1200
QY	1248	CTGAGCATTATAAGAGACCAAGATGAATGCTGCTGTAATGAATATCTCAGCTGGA	1307
DB	1201	CTGAGCATTATAAGAGACCAAGATGAATGCTGCTGTAATGAATATCTCAGCTGGA	1260
QY	1308	GGAGATGGCTCTCTATGCTGCTTCCATCACCCTGGGGAAGATTAAAGCATGACAAAA	1367
DB	1261	GGAGATGGCTCTCTATGCTGCTTCCATCACCCTGGGGAAGATTAAAGCATGACAAAA	1320
QY	1368	GAGAGGACAGATATTTCTCTCTAAGAGCTGATGCTTCAATTTCCATTTCTCAGCTC	1427
DB	1321	GAGAGGACAGATATTTCTCTCTAAGAGCTGATGCTTCAATTTCCATTTCTCAGCTC	1380
QY	1428	TCAGAAATTTGAGTTCATGAGATTTTGGCTCAGTTCCTTTTGGGAAAGTATATAAGGA	1487
DB	1381	TCAGAAATTTGAGTTCATGAGATTTTGGCTCAGTTCCTTTTGGGAAAGTATATAAGGA	1440
QY	1488	CGATGCAAGAAATAAATAGTGGCTATAAAAGTTATCGAGCAATACCTACTGCTCCAAG	1547
DB	1441	CGATGCAAGAAATAAATAGTGGCTATAAAAGTTATCGAGCAATACCTACTGCTCCAAG	1500
QY	1548	TCAGATGAGATATGTTTGGCGAGAGTGTCCATTTCTCTGCGAGCTCAATCATCCCTGC	1607
DB	1501	TCAGATGAGATATGTTTGGCGAGAGTGTCCATTTCTCTGCGAGCTCAATCATCCCTGC	1560
QY	1608	GTAATTTGAGTGTGGTGTGCTTGAATGATCCAGCCAGTTTGGCATTGTCACTCAA	1667
DB	1561	GTTGTTTCAATTTGTTGGTGTGCTGCTGATGACCCAGTCAATTTGCTCACTCAG	1620
QY	1668	TACATATCAGGGGTTCTCTGTTCTCCCTCTTCTATGAGCAAGAGATTTCTTGAATTTG	1727
DB	1621	TACATATCAGGGGTTCTCTGTTCTCCCTCTTCTATGAGCAAGAGATTTCTTGAATTTG	1680

QY 1728 CAGTCTAAATTAATTAATTCAGTAGATGTTGCCAAGGCATGGAGTACCTTCACAACTG 1787
DB 1681 CAGTCTAAATTAATTAATTCAGTAGATGTTGCCAAGGCATGGAGTACCTTCACAACTG 1740
QY 1788 ACACAGCCAAATATACATCGTACTTGAACAGTCACAAATATTTCTCTATGAGGATGG 1847
DB 1741 ACCCAGCCAAATATACATCGTACTTGAACAGTCACAAATATTTCTCTATGAGGATGG 1800
QY 1848 CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTCTACAGTCTCTGGATGAAGACAAC 1907
DB 1801 CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTCTCTGAGTCCCTGGATGAAGACAAC 1860
QY 1908 ATGACAAACAACTGGGAACCTCGTGTGGATGGCTCTGAGTGTGTTCACGAGTGCAC 1967
DB 1861 ATGACAAACAACTGGGAACCTCGTGTGGATGGCTCTGAGTGTGTTCACGAGTGCAC 1920
QY 1968 CGGTACACATCAAGCAGATGCTTTCAGCTATGCTCTGTGTGTGTGGGAAATTTCTCACT 2027
DB 1921 AGATACACATCAAGCAGTATGCTTTCAGTACTCCCTGTGTGTGGGAGTCTCTCACT 1980
QY 2028 GCGGAAATTCATTCGCTCACTCAAGCCAGCGGCTCGGCAGACAGATGGCTTACCCAC 2087
DB 1981 GGAGAAATTCATTCGCTCACTCAAGCCAGCGGCTCGGCAGACAGATGGGATATCAC 2040
QY 2088 CACATCAGACCTCCCATTTGGCTATTCATTCAGGAGCCATATCATCTCTGTGTATACGA 2147
DB 2041 CACATCAGACCTCCCATTTGGCTATTCATTCAGGAGCCATATCATCTCTGTGTATACGA 2100
QY 2148 GGGTGGAAACGATGCTCTGAAAGGAGACCCGAATTTTCTGAAGTGTGTCAAGATTAGAA 2207
DB 2101 GGTGAATGATGCTCTGAAAGGAGACCCGAAGTTCCTGAAGTGTGTCAAGATTAGAA 2160
QY 2208 GAGTGTCTCTGAAACATTTGAGTCTGCTGATGCTCTGATCAAGTACAGCAGTGGTCTCTC 2267
DB 2161 GAGTGTCTCTGAAACATTTGAGTCTGCTGATGCTCTGATCAAGTACAGCAGTGGTCTCTC 2220
QY 2268 TCACCT 2327
DB 2221 TCACCT 2280
QY 2328 GCAGCATTAAGAGTCTGTTTGAATATGCTCTTAAATGCAAGGCTCTCTATGCTGTCT 2387
DB 2281 GCAGCATTAAGAGTCTGTTTGAATATGCTCTTAAATGCAAGGCTCTCTATGCTGTCT 2340
QY 2388 TTGTCCTCAAGTCTGGAACATATTCCTCTCAAGTCTCTCTTGGAGGAGATGAAGA 2447
DB 2341 TGGTCCCAAGTCTGGAACATATTCCTCTCAAGTCTCTCTTGGAGGAGATGAAGA 2400
QY 2448 AGTCTTCAATACACACCACTTACAAATATGCTATGCTATGCTATGCTATGCTATGCT 2507
DB 2401 AGCACCAGTATTAACCTGTTGACAAATACGCTATGCTATGCTATGCTATGCTATGCT 2460
QY 2508 CATTTTCATTTCTGCGGAAATAGTAGAGCTTTTGGAGGAGATGAAGA 2552
DB 2461 CACCTTCACTCCCGCCAGACGACGCAACTTTTGGAGGAGATGAAGA 2505

RESULT 8

US-10-626-173-9
; Sequence 9, Application US/10626173
; Publication No. US20040110232A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaselan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT FILING DATE: US/10/626,173
; PRIOR APPLICATION NUMBER: 2003-07-24
; PRIOR APPLICATION NUMBER: US/09/947,199A
; PRIOR APPLICATION NUMBER: US 60/111,936
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14

; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (2505)
US-10-626-173-9

Query Match 62.6%; Score 1892.2; DB 17; Length 2505;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;
QY 48 ATGGGAAATTAATAATCTAGACCAACCCAAACTTGTACTGTATGAATGGAAGAAAAAGTC 107
DB 1 ATGGGGAATTAATAATCTAGACCAACCCAAACTTGTACTGTATGAATGGAAGAAAAAGTC 60
QY 108 AGTGAATCATATGTTATCATATAGAAAGATTAGAAAGTACCTGCAGATCAAGGAAAAA 167
DB 61 AGTGAATCTTACGCTATTATCATAGAAAGCTGGAGGATAACCTGCAGATCAAGGAAAAA 120
QY 168 GAACTGACAGAACTAAGGAATATATTTGGCTCTGTATGAAGCCCTTCAGTAAAGTCAATTTA 227
DB 121 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGTATGAAGCCCTTCAGTAAAGTCAATTTA 180
QY 228 AATTACCGCACTGAAATGGGCTGTCTACTTTCATTTAATGTTGCAATTTGTGGAGGCAAG 287
DB 181 AATTACCGCACTGAAATGGGCTGTCTACTTTCATTTAATGTTGCAATTTGTGGAGGCAAG 240
QY 288 AATACATATTTGAACTCTTATGTTGAAGGGCTCCGCCATCTCGACTGACAAAGAAAT 347
DB 241 AAGTCAATATCCGTGCGCTTATGTTAAAGGGCTCCGCCATCTCGACTGACAAAGAAAT 300
QY 348 GGATTTACAGCCTTTGCAATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG 407
DB 301 GGGTTTCAGCTCTGCACCTGGCCGTTTACAAGGACAGCCCGAACTTATCACTTCACTG 360
QY 408 CTTCAGTGGAGCTGATATACAGCAGTTGATAGCGTGGCTCACTGCCCTCCATATT 467
DB 361 TTGCACAGCGGACAGATGTTTCAGCAAGTGGGATAGCGTGGCTCACAGCCCTCCACATA 420
QY 468 GCTCAATAGCTGGCCACCTAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 527
DB 421 GCTCAATAGCTGGACACCCAGAGGCTGCAGAGGCTGCTACAAATGGGSCCAATGTG 480
QY 528 AATATCAAGATGAGTTTTTTTCACTCCATTTGCAATATTGCAAGCTATCTATGGACATGAA 587
DB 481 AATGTTCAAGATGCGCTCTTTCACCCCACTGCAATTTGCAGCCCTACTTATGGGACCGAG 540
QY 588 CAGTCACTCGCTCTCTTTTGAATTTGTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 647
DB 541 CAGTCACTCGCTCTCTTTTGAATTTGTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 648 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAATATTGCAAAATCTTTGATG 707
DB 601 GACAGGCTCTGCACCTGGGCTCTGCAAAAGGCTTCTTCAACATTTGTGAAATCTCGTGA 660
QY 708 GAAGAAGGCAAGCAAGCATGTAATGCTCAAGATAATGAAGACATGCTCCACTCCAT 767
DB 661 GAAGAAGGCAAGCAAGCATGTAATGCTCAAGATAATGAAGACATGCTCCACTCCAT 720
QY 768 TTCTGTTCTCGATTGGAACACCATGATAGTTTAAAGTATCTGCTCAAAAGTGAATTTGAA 827
DB 721 TTCTGTTCTCGATTGGAACACCATGATAGTTTAAAGTATCTGCTCAAAAGTGAATTTGAA 780
QY 828 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGCACTGCTACCAAT 887
DB 781 GTCCAGCCTCATCGTCAATTAACATCTATGTTGAGACACTCTTTTGACCTGCACTGCTACAA 840

```
QY 888 GGCAAAATTTGAAGTTCGAAGGAATCATCCAAATATCAAGAACGAAAGTCTGACTAAG 947
Db 841 GGAATATTTGAAGTTCGAAGGAATTTGCCAGGTAAACAGAACTCAAGTCTGACTAAG 900
QY 948 GAAAAATCTTCAAGTGAAGACAGCTTTTTCATAGTGTGTGATGTTGACATATGCGCAAGAGCAATTCAC 1007
Db 901 GAAAAATCTTCAAGTGAAGACAGCTTTTTCATAGTGTGTGATGTTGACATATGCGCAAGAGCAATTCAC 960
QY 1008 CTAGTCAAAATTTCTTCTTGATCAGAATGATATAAACAATCAACCAACAAAGAAAGGATGGG 1067
Db 961 CTGCTCAAAATTTCTTCTTGATCAGAATGATATAAACAATTAACCAACGAGGAGAGATGGG 1020
QY 1068 CACATGGGATTAACATCTGCTTCTACACAGGTGCATATTCGGCTGGTTCAGTTCCTTACTG 1127
Db 1021 CACAGGATTTGACATCTGCTTCTACACAGGCTATACCGCTGGTTCAGTTCCTTACTG 1080
QY 1128 GATAATGGAGCTGATATGAATCTAGTGGCTTTGTGATCCCAGCAGGTCTAGTGGTGAAGAAA 1187
Db 1081 GATAATGGTGCAGATATGAATCTTGTCCGTTGTGATCCCAGCAGGTCTAGTGGTGAAGAAA 1140
QY 1188 GATGCGCAGACATGTTGATGTGGGCTTATGAAAAGGGCATGATGCCATTTGTCAACATC 1247
Db 1141 GATGAGCAGACATGTTGATGTGGGCTTACGAGAAGGACATGATGCCATTTGTACATC 1200
QY 1248 CTGAAGCAATTAAGAGACACACAAGATGAATTCGCCCTGTAATGAATATTTCTCAGGCTGGA 1307
Db 1201 CTGAAGCACTTAAGAGACACCCAGGAGGAGTGCATGTAACGAATATTTCCCGAGCTGGA 1260
QY 1308 GAGAGATGGCTCCTATGATGCTGTTTCAATCAGCTTGGGGAAGATTAAGAGATGAAGAAA 1367
Db 1261 GAGAGATGGCTCCTATGATGCTGTTTCAATCAGCTTGGGGAAGATTAAGAGATGAAGAAA 1320
QY 1368 GAGAGGACAGATATTTCTCTCTCAAGAGTGGATGCTTCAATTTCCATCTTCAGCTC 1427
Db 1321 GAGAGGACAGATGTTCTCTCTCTGAGGCTGAATACCTCCCGCTTCATCTCCAACTC 1380
QY 1428 TCAGAAATTCAGTTCATGAGATATATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487
Db 1381 TCCGAAATCGAGTTCACAGAGAATATCGGCTCGGTTCTCTTTGGGAAAGTCTATAAGGG 1440
QY 1488 CGATGCAGAAATAAATAGTGGCTATAAAGCTTATCGAGGCAATACCTACTGCTCCAAAG 1547
Db 1441 CGATGCAGAAATAAATAGTGGGATCAACAGATCCGAGGCAACACCTACTGCTCCAAAG 1500
QY 1548 TCAGATGTGATATGTTTTCGCCAGAGGTGTCCATTTCTGTCAGCTCAATCATCCCTGC 1607
Db 1501 TCAGAGTGTGATATGTTTTCGCCAGAGGTGTCCATTTCTGTCGAGCTCAACACCCCTGC 1560
QY 1608 GTAAATTCAGTTTGGGTGCTTCTGCTTGAATGATCCAGCCAGTTTGCATTTGTCACTCAA 1667
Db 1561 GTGGTTTCAGTTTGGGTGCTTCTGCTTGAATGATCCAGCCAGTCAAGTTTGCATTTGTCACTCAG 1620
QY 1668 TACATATCAGGGGTTCTCTGTTTCTCCCTCTCATGAGCAGAGAGGATTTCTTGAATTG 1727
Db 1621 TACATTTACAGAGGCTCCCTGTTCTCCCTGCTTCATGAACAGAGAGATTTCTGACTTG 1680
QY 1728 CAGTCTAAATTAATATTCAGTAGATGTTTGCAAAGGATGAGTACCTTCACAACTTG 1787
Db 1681 CAGTCTAAATTAATATTCAGTGGTAGAGTTCGCAAGGCAATGAGTACCTGACAGCTTG 1740
QY 1788 ACACAGCCAAATTAACATCTGACTTGAACAGTCAATATTTCTCTCTATGAGATGGG 1847
Db 1741 ACCAGCCAAATTAACACCCGACCTGAACAGGCAATATTTCTGCTCTATGAGGATGGC 1800
QY 1848 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAAGACAAAC 1907
Db 1801 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTGAGTCCCTGGATGAAGACAAAC 1860
QY 1908 ATGACAAAAACAACCTGGGAACCTCCGTTGGATGGCTCCTGAGGTGTTTCAGCAGTGCCT 1967
Db 1861 ATGACAAAGCAGCCAGGGAACCTCCGCTGGATGGCCCTGAGGTGTTTCACAGTGCACG 1920
QY 1968 CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTCTGTGCTGTGGGAAATTTCTCACT 2027
```

```
Db 1921 AGATACACCATCAAGCTGATGTCTTTCAGTTHACTCCCTGTGTGTGGAGCTCTCTACT 1980
QY 2028 GGGCAATTTCCATTTCCGCTCATCTCAAGCCAGCGGTGGCGCAGCAGACATGGCTTACCAC 2087
Db 1981 GGGCAATTTCCATTTCCGCTCATCTCAAGCCAGCGGTGGCGCAGCAGATATGGCGTATCAC 2040
QY 2088 CACATCAGACCTTCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTCTGTATAGA 2147
Db 2041 CACATCAGACCCGCCATCGGCTATTCCATTTCCCAAGCCCATATCATCTCTCTGTATCGG 2100
QY 2148 GGGTGAAGCGCATGTCTCTGAAGGAAGACCCGAATTTTCTGAAGTTGTCTATGAAGTTAGAA 2207
Db 2101 GGGTGAAGTGCATGTCTCTGAAGGACGACAGAGTTCTCTGAAGTCTGTAGCAAACTGGAG 2160
QY 2208 GATGTCTCTGCAACATTTAGCTGATGTCTCTGATCAAGTAAACAGAGTGGGTCTCTC 2267
Db 2161 GAGTGGCTATGCAATTTGAGGCTCATGTCTCCAGCATCAAGTAAACAGCAGTGGCTCTCTG 2220
QY 2268 TCACCTTTCTTCTTCTTCTGATTTGCCGTGTAACCCGGGAGGACCTTGGCGGAGTCAATGTG 2327
Db 2221 TCACCTTTCTTCTTCTTCTGATTTGCCATTTGCCGTGAGCCGGGAGGCTTGGCCGAGCCACGTG 2280
QY 2328 GCAGCATTTAAGAAAGTGGTTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2387
Db 2281 GCAGCCTTACGGAGCCGTTTTCGATTTGAGTATGCCCTAAATGCAAGTCTCTATGCTGCTG 2340
QY 2388 TTGTCCCAAGTGTCTGCAATATTTCTCTCAAGTCTGTCTTTTGGAGGAGATGAAAGA 2447
Db 2341 TGTTCCCAAGTGTGGAACACACTCTTAATCCGGGCTGTCTTTTGGAGGAGATGAAATAGG 2400
QY 2448 AGTCTTTCAATACACACCCATTTGACAAATATGCTATGATCCGATCCCATGAGTCTCAATG 2507
Db 2401 AGCACCAGTATTTCAACTCTTGACAAATACGGCTATGTGTCTGATCCCATGAGCCTGACG 2460
QY 2508 CATTTTCATTTTCCGAAATAGTAGCAGCTTTTGGAGCAGCAGC 2552
Db 2461 CACCTTCACTCCCGCAAGACGACAGCAACTTTTGGAGCAGCAAC 2505
```

RESULT 9

```
US-09-833-381-1151
; Sequence 1151, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5900-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 1151
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(740)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1151
```

```
Query Match 14.1%; Score 427.6; DB 9; Length 740;
Best Local Similarity 78.4%; Pred. No. 1.7e-106;
Matches 519; Conservative 0; Mismatches 142; Indels 1; Gaps 1;
```

```
QY 16 CGSCCCTGGAGAGAGAAAGAACTTATAAATGGAATTTATAAATCTAGACCAACCC 75
Db 41 CAGCACAGGAGAGAAAGAAAGAACTTCTTTAAATGGGAATTTACAAATCCAGACCAAC 100
QY 76 AAACCTGTACTGATGAATGGAAGAAAAAGTCAGTGAATCATATGTTTATCACAATAGAAA 135
```

```
Db 101 AGACTTGTCTGATGAATGGAAGAAAGTTAGTGAATCTTAGCTATTATCATAGAAA 160
Qy 136 GATTAGAGATGACCTGCAGATCAAGGAAAGAAAGTACAGAACTAAGGAATATATTG 195
Db 161 GCGTGAAGATAACCTGCAGATCAAGAAATGATTTCAAGAACTAAGGCAATCTTTG 220
Qy 196 GCTCTGATGAAGCTTCAGTAAAGTCAATTTAAATACCGACCTGAAAATGGGCTGTCTC 255
Db 221 GCTCTGATGAAGCTTCAGTAAAGTCAATTTAAATACCGACCGTGGGCTGTCC 280
Qy 256 TACTTCAATTTATGTTGCAATTTGAGGAGCAAGAAATCACAATTTGAACTCTTATGTA 315
Db 281 TCGTACACCTCTGCTGTCTGTGCGCGCAACAAGTCAATATCCGTGCGCCCTTATGTTAA 340
Qy 316 AAGGCGCTCGGCCATCTCGACTGACAAAGAAATGGAATTTACAGCGCTTGCAATTTAGCAGTTT 375
Db 341 AAGGCGCTCGGCCATCTCGACTGACGAGAAATGGTTTCCAGCTCTGCACTGGCGGTTT 400
Qy 376 ACAAGGATAATGCAGAAATGATCACTTCTCTGCTTTCAC-AGTGGAGCTGATATACAGCAG 434
Db 401 ACAAGGACAGCGCGGAATTTATCACTTCACTGTTTGCAAGCGGAGCAGATGTTACAGCAA 460
Qy 435 GTTGATACGGTGGGCTCACTGCCCTCCATATTCCTACAATAGCTGCCACCTAGAGGCT 494
Db 461 GTGGATACGGTGGGCTCACAGCCCTCCACATAGCTGCAATAGCTGGACACCCAGAGGCT 520
Qy 495 GCTGATGCTGTGTTGCAACATGAGGCTAATGCTCAATATTCAGATGAGTTTTCAC 554
Db 521 GCANAAGTCTGNTGCAACATGGGCGCAACGCTGAATGTTCAAGATGCGGCTCTCTTCCACC 580
Qy 555 CCATTGCATATTGACGCTACTATGGACATGAACAGGTAACGCTCTCTTTTGAATTT 614
Db 581 CCACTGCACATTGACGCTACTATGGGCAAGCAGGTAACGAGTGTCTCTTTGAAGTTT 640
Qy 615 GGTGCTGATGTAATGTAAGTGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCA 674
Db 641 GGGGTGATGTCNATGTAAGCGGTAAAGTTGGGNACAGGCTCGCGCCCTGGCCNCCGNA 700
Qy 675 AA 676
Db 701 AA 702
```

RESULT 10

US-09-833-381-1149

; Sequence 1149, Application US/09833381

; Patent No. US20020132090A1

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1149

; LENGTH: 616

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-833-381-1149

Query Match 14.1%; Score 427.4; DB 9; Length 616;

Best Local Similarity 81.8%; Pred. No. 1.7e-106;

Matches 505; Conservative 0; Mismatches 111; Indels 1; Gaps 1;

```
Qy 338 GACAAGAAATGGAATTTACAGCTTGCATTTAGCAGTTTACAAGGATAATGCAGAAATGAT 397
Db 1 GACGAGAAATGGGTTTCCAGCTGTGCACCTGGCGGTTTACAAGGACAGCGCGGAATTTAT 60
Qy 398 CACTTCTCTGCTTCCAGTGGAGTGTATATACAGCAGGTTGGATACGGTGGGCTCACTGC 457
```

```
Db 61 CACTTCTCTGTTGCACAGCGGAGCAGATGTTTACGAAAGTGGGATACGGTGGGCTCACAGC 120
Qy 458 CCTCCATATTGCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTTTGCACATGG 517
Db 121 CCTCC-CATAGCTGCAATAGCTGACACCCAGAGGCTGCAGAAGTGTCTCTSCAACATGG 179
Qy 518 AGCTAATGTCAATATTCAAGATGCAAGTCTTTTCTCCTCCATTCGATTTGAGGCTACTA 577
Db 180 GGCCACAGCTGAATGTTTCAAGATGCCGTCTTCTTCACTCCACTGCAATTTGCAGGCTACTA 239
Qy 578 TGGACATGAACAGGTAACCTCGCTCTTTTGAATTTGCTGCTGATGTAATGTAAGTGG 637
Db 240 TGGCAGCAGCAGGTAACCAAGTGTCTTTTGAATTTGCTGCTGATGTAAGTGG 299
Qy 638 TGAAGTTGGAGATAGACCCCTCCACTAGCATCTGCAAAAGGATCTTGAATATTGCAAA 697
Db 300 TGAAGTTGGGACAGGCTCTGCACCTGCGCTCTGCAAAAGGCTCTTCAACATTTGTGAA 359
Qy 598 ACTTCTGATGGAAGAGGAGCAGCAAGATGTAATGCTCAAGATATGAAGACCATGT 757
Db 360 ACTCCTGTAAGAGAGGAGCAACAGATGTAACGCTCAGGACATGAAGACCATGT 419
Qy 758 CCCACTCCATTTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTCTCCAAAG 817
Db 420 CCCTCTGCACTTCTGTTCTCGATTTGGACACCAACATATAGTGAGCTACCTGCTCCAG 479
Qy 818 TGATTTGGAAGTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACACCTGCG 877
Db 480 TGACTTAGAGGTCGAGGCTCACGTCATTAACTATATGTTGACATCTCTTTCACCTGCG 539
Qy 878 ATGCTACAATGGCAATTTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAAACAGAAAG 937
Db 540 ATGCTACAATGGCAATTTTGAAGTTGCCAAGGAAATTTGCCAGGTAACAGGAACTGAAAG 599
Qy 938 TCTGACTAAGGAAACCA 954
Db 600 TCTGACTAAGGAAACCA 616
```

RESULT 11

US-09-833-381-1150

; Sequence 1150, Application US/09833381

; Patent No. US20020132090A1

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1150

; LENGTH: 736

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: misc feature

; LOCATION: (1), (736)

; OTHER INFORMATION: n = A, T, C or G

US-09-833-381-1150

Query Match 13.7%; Score 414.6; DB 9; Length 736;

Best Local Similarity 81.2%; Pred. No. 6.6e-103;

Matches 480; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

```
Qy 312 TTGAAGGGCTCGCCCATCTCGACTGCAAGAAATGGATTTACAGGCTTGCATTTAGCA 371
Db 1 TTAAGGGCTCGGCTCATCTCAGACTGACGAGAAATGGGTTTCCAGCTCTGACCTGGCC 60
Qy 372 GTTTACAAGGATAAGTACGAAATGATCACTTCTCTGTTTACAGTGGAGCTGATATACAG 431
```

Db 61 GTTTAAGGACGACCGGACATATACCTTCACTGTGACAGCGGACGAGTGTTCAG 120
Qy 432 CAGTTGGATACGGTGGCCCTCACTGCGCTCCATATGCTACATAGCTGGCCACCTAGAG 491
Db 121 CAACTGGGATACGGTGGCCCTCAGAGCCCTCCATAGCTGCAATAGCTGGACCCAGAG 180
Qy 492 GCTGCTGATGCTGTGTTGCAACATGAGGCTAATGTCAATATTCAAGATGAGTGTTC 551
Db 181 GCTGCAAGATGCTGTACAAATGGGCGCATGTGATGTTCAAGATGCGCTTCTTC 240
Qy 552 ACTCCATTGCATATGAGCGGTACTATGGAATGAAACAGGTAACTGCGCTTCTTTGAAA 611
Db 241 ACCCCACTGCACATGAGCGCTACTATGGGACAGGAGTAAACAGTGTCTCTTTGAAG 300
Qy 612 TTTGGTCTGCTGATTAAGTAAAGTGGTGAAGTGGAGATAGACCCCTCCACCTACATCT 671
Db 301 TTTGGTCTGATGTCAATGAGCGGTGAAGTGGGACAGGCTCTGACCTGGCTCT 360
Qy 672 GCAAAAGGATCTTTGATATTTGCAAACTCTTGATGGAAGAGGACAGCAAGCAGATGTG 731
Db 361 GCAAAAGGCTTCTTCAACATTGGAACCTCTGTTAGAAAGGAGCAAGCAGATGTG 420
Qy 732 AATGCTCAAGATTAAGACCAATGTCCTCCACTCCATTTCTGTTCTGATTTGGACACCAT 791
Db 421 AAGCTCAGACCAATGAAGACCAAGTCTCTGCTCTGCTCTGCTGATTTGGACACCAT 480
Qy 792 GATATAGTAAAGTATCTGCTGCAAGTGAATTTGGAAGTCAACCTCATGTTGTAATATC 851
Db 481 AATATAGTGAAGTACCTGCTCCAGATGACTTANAGTCCAGCTCACGTCAATTAACATC 540
Qy 852 TATGGAGATACCCCTTACACCTGGCATGTACAAATGGCAAAATTTGAAGTT 902
Db 541 TATGGTACACCTCTTTGCACTGGCATGTACAATGGAAATTTGAANTT 591

RESULT 12

US-09-833-381-1147/c
; Sequence 1147, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1e1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1147
; LENGTH: 304
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(304)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1147

Query Match 6.1%; Score 184.2; DB 9; Length 304;
Best Local Similarity 79.6%; Pred. No. 9.5e-40;
Matches 242; Conservative 0; Mismatches 59; Indels 3; Gaps 2;
Qy 313 TGAAGGCTCGGCCATCTCGACTGACAAAGATGATTAAGCC--TTGATTTAGC 370
Db 304 TAAAGGCTCGGCTCCATCCAGACTGACGAGAATGGTTTCAAGCTTTTGCACTGCC 245
Qy 371 AGTTTCAAGGATTAATGCAAGATTGATCAC-TTCTCTGCTTACAGTGAGCTGATATAC 429
Db 244 GTTTACAGGACAGCCGGATTTATCATTTTCACTGTTGACAGCGGAGCAGATGTC 185
Qy 430 AGCAGTTGGATACGGTGGCTCACTGCGCTCCATATGCTACAATAGCTGGCCACCTAG 489

Db 184 AGCAAGTGGGATACGGTGGCTCAGAGCCCTCCACATAGCTCAATAGCTGGACCCAG 125
Qy 490 AGGCTGCTGATGCTGCTTGGCAACATGGAGCTAATGTCAATATTCAAGATGAGTTTTT 549
Db 124 AGGCTGCAAGTGGCTGATGCAACATGGGCGCAAGTGAATGTTCAAGATGCGCTTCT 65
Qy 550 TCACCTCCATTGCATATTGAGCGGTACTATGGAACATGAACAGAGTAACTGCGCTTCTTTGA 609
Db 64 TCACCCACTGCACATTGCGGNTACTATGGGACAGGACAGGTAAACAGTGTCTTTGA 5
Qy 610 AATT 613
Db 4 AGTT 1

RESULT 13

US-10-424-599-53404
; Sequence 53404, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Tongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 53404
; LENGTH: 1953
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_19237C.1
US-10-424-599-53404

Query Match 3.6%; Score 109.6; DB 13; Length 1953;
Best Local Similarity 51.3%; Pred. No. 1.3e-18;
Matches 390; Conservative 0; Mismatches 349; Indels 21; Gaps 5;
Qy 1452 ATTGGCTCAGTTCTTTTGGGAAGTATATAAGGACGATGAGAAATAAATAGTGGCT 1511
Db 884 ATTGCATCTGGGTCATATGTGAACCTATTTAAAGGTGTATATTGAGCCAGGAAGTGGCC 943
Qy 1512 ATAAAGCTTATCGAGCCCAATACCTACTGCTCCAAGTCAGATGTGGATATGTTTCCCGA 1571
Db 944 ATCAAAGTTCTCAAGGCTGACCATGTAAACTCAGAATTCAGAGAGA---GTTCCACAG 1000
Qy 1572 GAGGTGCCATTCTCTGCCAGCTCAATCATCCCTGGTAAATTCAGTTTGTGGTGGCTTGC 1631
Db 1001 GAAGTCTTATATCATGAGAAGGTTGACACACAAGAATGTTGTACAATTCATAGAGCTGT 1060
Qy 1632 TTGAATGATCCAGCCAGTTTGGCATTGTCACTCAATACATATCAGGGGTTCTCTGTTC 1691
Db 1061 ACCAGTCCCCACGC---TTATGCATAGTAACAGAATTTATGCTGTTGGAAAGTGTGAT 1117
Qy 1692 TCCTCTCTTCATGAGCAGAGAGATCTTGTATTTGATTTGCACTCTAAATTAATTTAGCAGTA 1751
Db 1118 GACTACCTACATAAGCAGAGAGGCTTTTTTAAATTTCTTACACTGCTCAAAGTAGCAATT 1177
Qy 1752 GATGTTGCCAAAGGATGGAGTACCTTCAACCTGACACAGCAATATTATACATCGTGAC 1811
Db 1178 GATGTTTCCAAAGGAATGAATCTACTTGGACCA-----ACATAATATATCCATAGAGC 1231
Qy 1812 TTGAACAGTCACAATATTCTTCTATAGAGTGGGCATGCTGTGGTGGCAGATTTGA 1871
Db 1232 TTGAAGGCTCCCAACCTTTTGTATGGATGAATTTGACTGTAAAGTGGCTGCTGATTTGGG 1291
Qy 1872 GAATCAAGATTTCTACAGTCTCTGGATGAAGACAACATGACAAAACACCTGGACCTC 1931

Db 1292 GTTGTAGAGTTAA-----GCTCAATCTGGCGTCATGACTGCAGAACTGGAAACATAT 1345
QY 1932 CATTGATGGCTCTGAGGTTCACGGAGTGCACTGGGTACACCATCAAGCAGATGTC 1991
Db 1346 CGATGGATGGCTCTGAGGTATAGAAC-----CAAGCGGTATGATCAACAAGGCTGATGA 1402
QY 1992 TTCAGTATGCTCTGTCTGTGGGAATTCCTCACTGGCGAAATTCATTTCGCTCATCTC 2051
Db 1403 TTTAGTTTGGAAATGTTTATGGAGTTGTCTACCGGAAGCTTCCATACGGAATTTA 1462
QY 2052 AAGCCAGCGGTGGCGGACAGACATGGCTTTACCAACATCAGACCTCCATGGCTAT 2111
Db 1463 ACCCCCTACAGCGAGCTATAGGAGTGTTCAAAAGGTTTGGACCCACCATCCCGAAG 1522
QY 2112 TCCATTCACAGCCCATATCATCTCTGTATAGAGGTGGAGCGCATGCTCTGAGGA 2171
Db 1523 AACACTCATCCAAAGTTTGTGGAGCTTCTGAGAGGTCTTGGCAGCAAGATCCTATATG 1582
QY 2172 AGACCCGAATTTCTGAAGTTGTATGAAGTTAGAAGT 2211
Db 1583 AGACCTGATTTCTCGAAATATCGAGATCTCGCAGCAGT 1622

RESULT 14

US-10-424-599-129156
; Sequence 129156, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 129156
; LENGTH: 3238
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MBT3847_87633C.1
US-10-424-599-129156

Query Match 3.6%; Score 107.4; DB 13; Length 3238;
Best Local Similarity 52.4%; Pred. No. 7.4e-18;
Matches 387; Conservative 0; Mismatches 331; Indels 21; Gaps 6;
QY 1453 TTGGCTCAGGTTCTTTGGGAAAGTATATAAGACCATGCAGAAATAAATAGTGCTA 1512
Db 1615 TTGGCTCTGGGTCATTTGGTGACTGTGACAGAGGTACATATTGTAGTCAAGATGTGGCTA 1674
QY 1513 TAAAAAGTTATCGAGCCAAATACCTACTGCTCCAGTCAGATGTGGATATGTTTGGCGAG 1572
Db 1675 TCAAA-GTCTTAAGCTTGAGCGCATAGTACAGATATGCTGAGG--GAGTTTGCACAGG 1731
QY 1573 AGGTGTCATTCCTGTCAGCTCAATCATCCCTCGGTAATTCAGTTTGGGGTGTGCT 1632
Db 1732 AAGTTTATATCATGAGGAAGATTCGACACAGAAATGTTGTTCAGTTCAITGGGCGCATG-- 1789
QY 1633 TGAATGATCCAGCCAGTTTCCCATTTGTCATCAATACATATCAGGGGGTCTCTGTCTCT 1692
Db 1790 -TACTAGGCCCCCAATCTTTGTTATGTTACTGAGTTTATGTTCTAGGGGAGCTTATATG 1848
QY 1693 CCTCTCTCATGACAGAAAGGATTCCTTGATTTTGCAGCTTAAATTAATTTATTCAGTAG 1752
Db 1849 ACTTTCTGCACAAACAAAGAGGTGTTTAAAGCTTCCATCTTTGCTAAAAGTAGCAATTG 1908
QY 1753 ATGTTGCCAAAGGCTAGGAGTACCTTCACACCTGACACGCCAATTTATACATCGTGACT 1812
Db 1909 ATGTTCCAGGAATGAATATTGTCACCA-----AAATATATTAATTCAGGGACC 1962

QY 1813 TGAACAGTCACATATTTCTTCTATGAGGATGGCATGCTGTGGTGGCAGATTTTGGAG 1872
Db 1963 TCAAGACTGCCATCTTCTGATGGATGAAAATGAAGTGTCAAGTTGTGTTTGGG 2022
QY 1873 AATCAAGATTTTACAGTCTCTGGATGAAGACAAATGACAAACAACTGGGAACCTCC 1932
Db 2023 TTGCAGGGTGCAA-----ACTCAGTCTGGAGTGATGACAGCTGAAACTGGAAACATACC 2076
QY 1933 GTTGGATGGTCTCTGAGGTGTTTCCAGCAGTGCACTGGGTACACCATCAAGCAGATGCT 1992
Db 2077 GTTGGATGGTCTCTGAGG---TCAATTGAACAAACATATGACCAAGGAGGAGATGTT 2133
QY 1993 TCAGTATGCTCTGTGTGTGGGAAATTTCTCACTGGCGAAATTTCCATTCGCTCATCTCA 2052
Db 2134 TCAGTTTGGAAATAGCTCTTTGGAGCTTTTAACTGGAGAACTGCTTACTCTTCCCTGA 2193
QY 2053 AGCCAGCGGTGGCGGACAGACATGGCTTACCAACATCAGACCTCCATTTGGCTATT 2112
Db 2194 CCCAATTACAAGCAGCAGTTGGCGTGGTGCAGAAAGGCTTGGCGCTTCAATTTCCAAA 2253
QY 2113 CCATTCCAAAGCCCATATCATCTCTGTATACGAGGTGGAAACGATGCTCTGAAGGAA 2172
Db 2254 ATACGACCCCAAGACTTTCTGAACTTCTTCAGCGGTGCTGGCAACAAGATCCACGCAA 2313
QY 2173 GACCCGAATTTCTGAAGT 2191
Db 2314 GACCAAAATTTCTCTGAAAT 2332

RESULT 15

US-10-425-114-22955
; Sequence 22955, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Skreen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22955
; LENGTH: 2207
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3592-081-E1_FLI
US-10-425-114-22955

Query Match 3.3%; Score 100.2; DB 13; Length 2207;
Best Local Similarity 52.1%; Pred. No. 5.4e-16;
Matches 328; Conservative 0; Mismatches 283; Indels 18; Gaps 4;
QY 1563 TTTTGGCAGAGGTGTCATTTCTCTGCCAGCTCAATCATCTCCGTTAATTCAGTTTGTG 1622
Db 1133 TTTGCTCAGGAAGTATATATTATGAGGAAGTCCGTCACAGAAATGTTGTGCAATTTATT 1192
QY 1623 GGTCTTGTGTTGAATGATCCAGCGAGTTTGCAATGTGTCATCAATATACATATCAGGGGT 1682
Db 1193 GGGCTTGCACTAAACCCCAACCTATGT--ATAGTACAGAAATTTATGTCGCTGGA 1249
QY 1683 TCTCTGTTCTCCCTCTTCATGAGCAGAGAGATTTCTTGATTTGAGTCTTAATTAATT 1742
Db 1250 AGTGTGATGATTACCTCCATAAACAAGAGTGTGTTTCAAACTTCTCTGTTAGTGA 1309
QY 1743 ATTGCACTAGATGTTGCCAAGGCATGGAGTACCTTCACAACCTGACACAGCAATATA 1802

```

Db 1310 GTTCCAAACGGATGTCTCAAAAGCATGAGCTACTTACACAGAAATAAT-----ATTATT 1363
QY 1803 CATCGTGAAGTGAACAGTACAAATATTCTCTATGAGGATGGGCATGCTGTGGTGGCA 1862
Db 1364 CATCGTGAATTTGAACATGCAACCTTCTGATGATGAATGGGACTCTTAAAGTTGCT 1423
QY 1863 GATTTTGGAGAAACAAGATTTCTACAGTCTCTGGATGAAGACAACATGACAAAAACAACCT 1922
Db 1424 GATTTTGGTGTGTGCACGTGTTAAAGCTCAATCTGGAGTT-----ATGACTGCAGAAACT 1477
QY 1923 GGGAACTTCGTTGGATGGGCTCCTGAGGTGTTCAACGAGTGCATCGGTACACCACTCAA 1982
Db 1478 GGTACTTACCGTTGGATGGGCCAGAGGTCAATTGACA---CAAACCTATGATCACACAG 1534
QY 1983 GCAGATCTCTTCAAGCTATGCTCTGTGTCTGTGGAAATTTCTACTGGCGAAATTCATTTC 2042
Db 1535 GCTGATGTTTTTATTGTTTGGAAATTTTGTGTGGAACTACTACAGGGAAGATTCCTTAT 1594
QY 2043 GCTCATCTCAAGCCAGCGGCTGGGCAGCAGACATGGCTTACCACCATCAGACTGCC 2102
Db 1595 GAGTACCTGACTCCACTACAAAGCGGTGTTGGTGGTGCAGAAAGGATTACGGGCTACA 1654
QY 2103 ATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGTGATACGAGGTGGAAACGCATGT 2162
Db 1655 ATTCCAAAGCATACTCAGCNAAGCTTTCGAGCTTCTTCAGAAATGCTGGCAGCAGGAC 1714
QY 2163 CCTGAAGGAAGACCCGAATTTTCTGAAGT 2191
Db 1715 CCTACCCAAAGACCAGACTTCTCCGAAAT 1743

```

Search completed: September 6, 2004, 11:52:06
 Job time : 922.158 secs

This Page Blank (uspto)

CC function. They also act as modulators of cellular differentiation and
CC cell death. Nucleic acids, or their fragments encoding CARP polypeptides
CC are useful for detecting CARP nucleic acids especially mRNA, in a sample.
CC CARP is useful for treating disorders associated with upregulation or
CC downregulation of cellular proliferation such as, cardiovascular
CC disorders (heart failure and hypertension) and disorders associated with
CC cell differentiation such as cancer and sarcoma

XX Sequence 3025 BP; 882 A; 622 C; 655 G; 866 T; 0 U; 0 Other;

Query Match		100.0%; Score 3025; DB 3; Length 3025;
Best Local Similarity		100.0%; Pred. No. 0;
Matches 3025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	GTGACCCACGCGTCGGCCCTGGAGAAAGGAGAACTTATAATAATGGGAAATTATA 60
DB	1	GTGACCCACGCGTCGGCCCTGGAGAAAGGAGAACTTATAATAATGGGAAATTATA 60
QY	61	AATCTAGACCAACCAAACTTGTACTGATGAATGGAAGAAAAAGTCAGTGAATCATATG 120
DB	61	AATCTAGACCAACCAAACTTGTACTGATGAATGGAAGAAAAAGTCAGTGAATCATATG 120
QY	121	TTATCACAATAGAAAGATTAGAGATGACCTGCGAGTCAAGGAAAAAGAACTGACAGAAC 180
DB	121	TTATCACAATAGAAAGATTAGAGATGACCTGCGAGTCAAGGAAAAAGAACTGACAGAAC 180
QY	181	TAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTAACCGCACTG 240
DB	181	TAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTAACCGCACTG 240
QY	241	AAAAATGGGCTGCTCTACTTCAATTTATGTTGCAATTTGTTGGAGGCAAGAAATCACATATTC 300
DB	241	AAAAATGGGCTGCTCTACTTCAATTTATGTTGCAATTTGTTGGAGGCAAGAAATCACATATTC 300
QY	301	GAACCTTTATGTTGAAAGGCTCCGCCATCTGACCTGACAGCAAGAAATGGAATTTACAGCCT 360
DB	301	GAACCTTTATGTTGAAAGGCTCCGCCATCTGACCTGACAGCAAGAAATGGAATTTACAGCCT 360
QY	361	TGCAATTTAGCAGTTTACAGGATTAATGCAGAAATGATCACTTCTCTGCTTCAAGTGGAG 420
DB	361	TGCAATTTAGCAGTTTACAGGATTAATGCAGAAATGATCACTTCTCTGCTTCAAGTGGAG 420
QY	421	CTGATATACAGAGTTGGATACGGTGGCCTCACTGCCCTCCATATTTGCTAGCAATAGCTG 480
DB	421	CTGATATACAGAGTTGGATACGGTGGCCTCACTGCCCTCCATATTTGCTAGCAATAGCTG 480
QY	481	GCCACTAGAGGCTGCTGATGCTGTTGCAACATGGAGCTTAATGTCAATATTTCAAGATG 540
DB	481	GCCACTAGAGGCTGCTGATGCTGTTGCAACATGGAGCTTAATGTCAATATTTCAAGATG 540
QY	541	CAGTTTTTTTCACTCCATTCATATTCAGGCTACTATGCAATGAAGCAAGGTAACCTCGCC 600
DB	541	CAGTTTTTTTCACTCCATTCATATTCAGGCTACTATGCAATGAAGCAAGGTAACCTCGCC 600
QY	601	TTCTTTGCAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGAGATAGACCCCTCC 660
DB	601	TTCTTTGCAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGAGATAGACCCCTCC 660
QY	661	ACCTAGCATCTGCAAAAGGATTTCTTGAATATTTGCAAACTCTTGTGGAAGGAGGAGCA 720
DB	661	ACCTAGCATCTGCAAAAGGATTTCTTGAATATTTGCAAACTCTTGTGGAAGGAGGAGCA 720
QY	721	AAGCAGATGTAATGCTCAGATTAATGAAGCCTATGTCCTCCATTTCTGTTCTCGAT 780
DB	721	AAGCAGATGTAATGCTCAGATTAATGAAGCCTATGTCCTCCATTTCTGTTCTCGAT 780
QY	781	TTGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACCTCATG 840
DB	781	TTGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACCTCATG 840
QY	841	TTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAATGGCAATTTGAAG 900
DB	841	TTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAATGGCAATTTGAAG 900

QY	901	TTGCCAAGGAATCATCCAAATATCAGAAACAGAAAGTCTGACTAAAGGAAAAACATCTTCA 960
DB	901	TTGCCAAGGAATCATCCAAATATCAGAAACAGAAAGTCTGACTAAAGGAAAAACATCTTCA 960
QY	961	GTGAAACAGCTTTTCATAGTGTCTGTACCTTATGCCAGAGCATTCACCTAGTCAAAATTC 1020
DB	961	GTGAAACAGCTTTTCATAGTGTCTGTACCTTATGCCAGAGCATTCACCTAGTCAAAATTC 1020
QY	1021	TTCTTGATCAGATGTCTATAAACATCAACCAAGGAGGAGGATGGGCACACTGGATTAC 1080
DB	1021	TTCTTGATCAGATGTCTATAAACATCAACCAAGGAGGAGGATGGGCACACTGGATTAC 1080
QY	1081	ACTCTGCTTGTCTACACGGTCAATTCGCCCTGCTTTCAGTTCTTACTTGGATTAATGGAGCTG 1140
DB	1081	ACTCTGCTTGTCTACACGGTCAATTCGCCCTGCTTTCAGTTCTTACTTGGATTAATGGAGCTG 1140
QY	1141	ATATGAATCTAGTGGCTTGTGATCCACAGAGGTCTAGTGGTGAAGGATGAGCAGACAT 1200
DB	1141	ATATGAATCTAGTGGCTTGTGATCCACAGAGGTCTAGTGGTGAAGGATGAGCAGACAT 1200
QY	1201	GTTTGATGTGGGCTTATGAAAAAGGCGATGATGCCATTTGTCACTCTCTGAAAGCATTTATA 1260
DB	1201	GTTTGATGTGGGCTTATGAAAAAGGCGATGATGCCATTTGTCACTCTCTGAAAGCATTTATA 1260
QY	1261	AGAGACCACAGATGAAATTCGCCCTGTAATGAATTTCTCAGCTGGAGGAGATGGCTCT 1320
DB	1261	AGAGACCACAGATGAAATTCGCCCTGTAATGAATTTCTCAGCTGGAGGAGATGGCTCT 1320
QY	1321	ATGTGCTGTGTCCATCACTCCCTTGGGGAAGATTAAAAAGCATGACAAAAGAGAGGAGGAGATA 1380
DB	1321	ATGTGCTGTGTCCATCACTCCCTTGGGGAAGATTAAAAAGCATGACAAAAGAGAGGAGGAGATA 1380
QY	1381	TTCTCTCTCTTAAGAGCTGGATTCCTTCACTTCCATCTTCCAGCTCTCAGAAATGAGT 1440
DB	1381	TTCTCTCTCTTAAGAGCTGGATTCCTTCACTTCCATCTTCCAGCTCTCAGAAATGAGT 1440
QY	1441	TCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGAGCATGACGAAATA 1500
DB	1441	TCCATGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGAGCATGACGAAATA 1500
QY	1501	AAATAGTGGCTTATAAAGCTTATCGAGCAATACCTACTGCTCCAGTCAAGTGGAGATA 1560
DB	1501	AAATAGTGGCTTATAAAGCTTATCGAGCAATACCTACTGCTCCAGTCAAGTGGAGATA 1560
QY	1561	TGTTTTGCCAGAGGTTCCATTTCTGCGAGCTCAATCATCTCCCTGCGTAAATTCAGTTTG 1620
DB	1561	TGTTTTGCCAGAGGTTCCATTTCTGCGAGCTCAATCATCTCCCTGCGTAAATTCAGTTTG 1620
QY	1621	TGGTGTCTGCTTGAATGATCCAGCGATTTGCCATTTGCTCACTCAATATATATATATAT 1680
DB	1621	TGGTGTCTGCTTGAATGATCCAGCGATTTGCCATTTGCTCACTCAATATATATATATAT 1680
QY	1681	GTTCTCTGTTCTCCCTCTTCTATGACGAGAGGATTTCTGATTTGAGTCAATTAATTA 1740
DB	1681	GTTCTCTGTTCTCCCTCTTCTATGACGAGAGGATTTCTGATTTGAGTCAATTAATTA 1740
QY	1741	TTATTGAGTATGTTTGGCAAGGATGAGTACCTTCAACCTGACACAGGCAATTA 1800
DB	1741	TTATTGAGTATGTTTGGCAAGGATGAGTACCTTCAACCTGACACAGGCAATTA 1800
QY	1801	TACATGCTGATTTGAACAGTCAATTTCTTCTATGAGGATGGGATGCTGTTGG 1860
DB	1801	TACATGCTGATTTGAACAGTCAATTTCTTCTATGAGGATGGGATGCTGTTGG 1860
QY	1861	CAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAACATGACAAACAC 1920
DB	1861	CAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAACATGACAAACAC 1920
QY	1921	CTGGGAACCTCGTTGGATGGCTCTGAGGTTTCAACAGTGGATCTGGTACACCATCA 1980
DB	1921	CTGGGAACCTCGTTGGATGGCTCTGAGGTTTCAACAGTGGATCTGGTACACCATCA 1980

QY 1981 AAGCAGATGCTTTCAGCTATGCTCTGTGCTGTGGGAAATTCCTCACTGCGGAAATTCAT 2040
Db AAF44702
XX AAF44702 standard; cDNA; 2508 BP.
AC AAF44702;
XX
XX
DT 27-MAR-2001 (first entry)
XX
DE Novel protein kinase cDNA, SEQ ID NO: 83.
XX
XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
XX
XX Homo sapiens.
OS
XX WO200073469-A2.
XX
XX 07-DEC-2000.
XX
XX 26-MAY-2000; 2000WO-US014842.
XX
XX 28-MAY-1999; 99US-0136503P.
XX
XX (SUGE-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
XX WPI: 2001-032161/04.
XX
XX P-PSDB; AAB65674.
PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers.
PS Disclosure; Fig 2; 310pp; English.
XX
CC The present sequence encodes a novel protein kinase. The nucleic acids
CC and the protein kinases they encode may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure, oxidative-
CC stress related disorders, chronic inflammatory bowel disease, chronic
CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,
CC psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive
CC disorders
XX
SQ Sequence 2508 BP; 722 A; 532 C; 555 G; 699 T; 0 U; 0 Other;

QY 48 ATGGGAAATATATAATCTAGACCAACCCAACTTGCTACTGATGAATGGAGAAAAAGTC 107
Db 1 ATGGGAAATATATAATCTAGACCAACCCAACTTGCTACTGATGAATGGAGAAAAAGTC 60

QY 108 AGTGAATCATATCTTATCATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGAAAA 167
Db 61 AGTGAATCATATCTTATCATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGAAAA 120

QY 168 GAACCTGACAGAACTAAGGAATATATTGGCTCTGATGAGCCTTCAGTAAAGTCAATT 227

Query Match 82.9%; Score 2508; DB 4; Length 2508;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 121 GAACTGACAGAACTAAGGAAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180
QY 228 AATTACCAGCTGAAAATGGGCTGTCTCTACTTCAATTTATGTTGCAATTTGTGGAGCAAG 287
Db 181 AATTACCAGCTGAAAATGGGCTGTCTCTACTTCAATTTATGTTGCAATTTGTGGAGCAAG 240
QY 288 AATACACATATTCGAACCTCTTATGTTGAAGGGCTCGCCCATCTGACATGACAAGAAAT 347
Db 241 AATACACATATTCGAACCTCTTATGTTGAAGGGCTCGCCCATCTGACATGACAAGAAAT 300
QY 348 GGAATTTACAGCCTTCGATTTAGCAGTTTACAGGATAATGCAAGATAATGCAAGATTTGATCACTTCTCTG 407
Db 301 GGAATTTACAGCCTTCGATTTAGCAGTTTACAGGATAATGCAAGATAATGCAAGATTTGATCACTTCTCTG 360
QY 408 CTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTGCGCTCCATATTT 467
Db 361 CTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTGCGCTCCATATTT 420
QY 468 GCTACAATAGCTGGCCACTAGAGCTGCTGATGCTGTTGCAACATGAGCTAATGTC 527
Db 421 GCTACAATAGCTGGCCACTAGAGCTGCTGATGCTGTTGCAACATGAGCTAATGTC 480
QY 528 AATATTCAGATGCACTTTTTCACATTCATTTGATATGAGAGCTACTATGACATGAA 587
Db 481 AATATTCAGATGCACTTTTTCACATTCATTTGATATGAGAGCTACTATGACATGAA 540
QY 588 CAGGTAACTCGGCTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA 647
Db 541 CAGGTAACTCGGCTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA 600
QY 648 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTTGAATATGCAAAACTCTTTGATG 707
Db 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTTGAATATGCAAAACTCTTTGATG 660
QY 708 GAAGAGGAGCAAGAGATGTGMAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 767
Db 661 GAAGAGGAGCAAGAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720
QY 768 TTCTGTCTCGATTTGGACACCATGATAGTAAAGTATCTGTCGCAAGTATTTGGAA 827
Db 721 TTCTGTCTCGATTTGGACACCATGATAGTAAAGTATCTGTCGCAAGTATTTGGAA 780
QY 828 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGGCATGCTACAAT 887
Db 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACACCTGGCATGCTACAAT 840
QY 888 GGCRAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947
Db 841 GGCRAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
QY 948 GAAACATCTTCAGTGAACAGCTTTTTCATAGTCTTTGATAGTCTGATGCAAGAGCATGAC 1007
Db 901 GAAACATCTTCAGTGAACAGCTTTTTCATAGTCTTTGATAGTCTGATGCAAGAGCATGAC 960
QY 1008 CTAGTCAAAATTTCTTTGATCAGNATGCTATAAATCATCAACCAAGAGGATGGG 1067
Db 961 CTAGTCAAAATTTCTTTGATCAGNATGCTATAAATCATCAACCAAGAGGATGGG 1020
QY 1068 CACATCGGATTAACCTCTGCTTGTACCCAGGTCACATTCGCTGGTTCAAGTCTTACTG 1127
Db 1021 CACATCGGATTAACCTCTGCTTGTACCCAGGTCACATTCGCTGGTTCAAGTCTTACTG 1080
QY 1128 GATAATGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGCTAGTGGTGAATA 1187
Db 1081 GATAATGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGCTAGTGGTGAATA 1140
QY 1188 GATGAGCAGATGTTTTCATGTTGGCTTATGAAAAGGGCATGATGCAATTTGTCACATC 1247
Db 1141 GATGAGCAGATGTTTTCATGTTGGCTTATGAAAAGGGCATGATGCAATTTGTCACATC 1200
QY 1248 CTGAAGCATTTAAGAGACCAAGATGAATGTCCTGTGAATGAATTTCTAGCCTGGA 1307

Db 1201 CTGAAGCATTTAAGAGACCAAGATGAATGTCCTGTATGAATATTTCTCAGCCTGGA 1260
QY 1308 GGAGATGGCTCCTATGATGTCGTCTGTCATCACCCCTGGGGAAGATTTAAAGCATGCAAAA 1367
Db 1261 GGAGATGGCTCCTATGATGTCGTCTGTCATCACCCCTGGGGAAGATTTAAAGCATGCAAAA 1320
QY 1368 GAGAAGCAGATATTTCTCCTCCTAAAGAGTGGATTCCTTCAATTTCCATCTTCAGCTC 1427
Db 1321 GAGAAGCAGATATTTCTCCTCCTAAAGAGTGGATTCCTTCAATTTCCATCTTCAGCTC 1380
QY 1428 TCAGAAATTCAGTTCCTAAGAGATTTATGCTCAGGTTCTTTTGGGAAGATATATAAGGA 1487
Db 1381 TCAGAAATTCAGTTCCTAAGAGATTTATGCTCAGGTTCTTTTGGGAAGATATATAAGGA 1440
QY 1488 CGATGAGAAATAAATAGTGGCTATAAAGAGTTATCGAGCCAAATACCTACTGCTCCCAAG 1547
Db 1441 CGATGAGAAATAAATAGTGGCTATAAAGAGTTATCGAGCCAAATACCTACTGCTCCCAAG 1500
QY 1548 TCAGATGAGATATGTTTTCGAGAGGTTCCATTTCTGCGAGCTCAATCATCCCTGC 1607
Db 1501 TCAGATGAGATATGTTTTCGAGAGGTTCCATTTCTGCGAGCTCAATCATCCCTGC 1560
QY 1608 GTAATTCAGTTCGAGGTTGCTTGAATGATCCAGCCAGTTTGGCATTGCTCACTCAA 1667
Db 1561 GTAATTCAGTTCGAGGTTGCTTGAATGATCCAGCCAGTTTGGCATTGCTCACTCAA 1620
QY 1668 TACATATCAGGGGTTCTCTGTTCTCCTCTTCATGAGCAGAGAGGATTTCTGATTTG 1727
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCTCTTCATGAGCAGAGAGGATTTCTGATTTG 1680
QY 1728 CAGTCTAAATTAATATTCAGTAGATGTTGCCAAGGATGGAGTACCTTCACAACTG 1787
Db 1681 CAGTCTAAATTAATATTCAGTAGATGTTGCCAAGGATGGAGTACCTTCACAACTG 1740
QY 1788 ACAGAGCAATATATACATCGTACTTGAACAGTACAAATATTTCTCTCTATCAGGATGGG 1847
Db 1741 ACAGAGCAATATATACATCGTACTTGAACAGTACAAATATTTCTCTCTATCAGGATGGG 1800
QY 1848 CATCTGTGTGGCAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAAC 1907
Db 1801 CATCTGTGTGGCAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAAC 1860
QY 1908 ATGACAAACCAACCTGGACCTCCGTTGATGCTGAGTGTTCAGCAGTGCACCT 1967
Db 1861 ATGACAAACCAACCTGGACCTCCGTTGATGCTGAGTGTTCAGCAGTGCACCT 1920
QY 1968 CGGTACACCATCAAGACAGATGCTTTCACTATGCTCTGTCTGTGCGAAATTTCTCACT 2027
Db 1921 CGGTACACCATCAAGACAGATGCTTTCACTATGCTCTGTCTGTGCGAAATTTCTCACT 1980
QY 2028 GGCRAATTTCAATTCGCTCATCTCAAGCCAGGCTCGGCGAGCAGACATGCTTACCAC 2087
Db 1981 GGCRAATTTCAATTCGCTCATCTCAAGCCAGGCTCGGCGAGCAGACATGCTTACCAC 2040
QY 2088 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCATATCATCTCTGCTGATAGGA 2147
Db 2041 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCATATCATCTCTGCTGATAGGA 2100
QY 2148 GGGTGGAAACGATGCTCTGAAGGAAGACCCGAATTTTCTGGAAGTGTCTATGAAGTAGAA 2207
Db 2101 GGGTGGAAACGATGCTCTGAAGGAAGACCCGAATTTTCTGGAAGTGTCTATGAAGTAGAA 2160
QY 2208 GAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGTCATCAAGTAAACAGCAGTGGGCTCTCTC 2267
Db 2161 GAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGTCATCAAGTAAACAGCAGTGGGCTCTCTC 2220
QY 2268 TCACCTTCTTCTTCTGATTCCTTGGTGAACCCGGGAGGACCTTGGCCGAGTCTATGTC 2327
Db 2221 TCACCTTCTTCTTCTGATTCCTTGGTGAACCCGGGAGGACCTTGGCCGAGTCTATGTC 2280
QY 2328 GCACATTAAGAGTCTGTTTTCGAATGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2387
Db 2281 GCACATTAAGAGTCTGTTTTCGAATGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2340

QY 2388 TTGTCCCAAAGTGTGACAAATATTCCTCTCAAGGTCTGTCTTTGGAGAGATGAAAGA 2447
DB 2341 TTGTCCCAAAGTGTGACAAATATTCCTCTCAAGGTCTGTCTTTGGAGAGATGAAAGA 2400
QY 2448 AGTCTTCAATACACACCCATTTGCAATATGATGGCTATGATFCCGATCCCATAGCTCAATG 2507
DB 2401 AGTCTTCAATACACACCCATTTGCAATATGATGGCTATGATFCCGATCCCATAGCTCAATG 2460
QY 2508 CATTTTCATTTCTTGGCCGAAATAGTAGCAGCTTTTGAGGACAGCAGCTGA 2555
DB 2461 CATTTTCATTTCTTGGCCGAAATAGTAGCAGCTTTTGAGGACAGCAGCTGA 2508

RESULT 4
AAAA47607
ID AAA47607 standard; cDNA; 2505 BP.
XX
AC AAA47607;
XX
DT 20-OCT-2000 (first entry)
XX
DE Human CARK (Cardiac related Ankyrin-Repeat Protein Kinase) cDNA.
XX
KW Cardiac related ankyrin repeat protein kinase; CARK; cytoskeleton;
KW cardiac cell growth factor receptor; cell differentiation; modulator;
KW regulator; detection; cellular proliferation; cardiovascular disorder;
KW heart failure; hypertension; cancer; sarcoma; ds.
XX
OS Homo sapiens.

XX
PH Key Location/Qualifiers
FT CDS 1..2505
FT /*tag= a
FT /product= "Human CARK"
XX
PN WO200034330-A1.
XX
PD 15-JUN-2000.
XX
PF 10-DEC-1999; 99WO-US029465.
XX
PR 11-DEC-1998; 98US-0111938P.
PR 14-APR-1999; 99US-00291839.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
XX Raju J;
XX
XX WPI; 2000-431275/37.
XX P-PSDB; AAB01470.
XX

XX New polynucleotide encoding cardiac-related ankyrin-repeat protein
XX kinase, useful for treating disorders such as cardiovascular disorders,
XX e.g. heart failure and cell differentiation disorders, e.g. cancer.
XX
XX Claim 1; Page 143-147; 161pp; English.
XX
XX CARK polypeptides are regulators of signal transmission from cellular
XX receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton
XX function. They also act as modulators of cellular differentiation and
XX cell death. Nucleic acids, or their fragments encoding CARK polypeptides
XX are useful for detecting CARK nucleic acids especially mRNA, in a sample.
XX CARK is useful for treating disorders associated with upregulation or
XX downregulation of cellular proliferation such as, cardiovascular
XX disorders (heart failure and hypertension) and disorders associated with
XX cell differentiation such as cancer and sarcoma

XX Sequence 2505 BP; 721 A; 532 C; 554 G; 698 T; 0 U; 0 Other;
XX
XX Query Match 82.8%; Score 2505; DB 3; Length 2505;
XX Best Local Similarity 100.0%; Pred No. 0;
XX Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 ATGGGAAATTATTAATCTAGACCAACCCAAACTTGTACTGTAGTAATGGAAGAAAAAGTC 107
DB 1 ATGGGAAATTATTAATCTAGACCAACCCAAACTTGTACTGTAGTAATGGAAGAAAAAGTC 60
QY 108 AGTGAATCATATGTTATCAACAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 167
DB 61 AGTGAATCATATGTTATCAACAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 120
QY 168 GAACTGACAGAACTAAGGAATATATTGGCTCTGTATGAAGCCCTTCAGTAAAGTCAATTTA 227
DB 121 GAATGACAGAACTAAGGAATATATTGGCTCTGTATGAAGCCCTTCAGTAAAGTCAATTTA 180
QY 228 AATTACCGCACTGAAAAATGGGCTGTCTCTACTTATTTATTTTGTGCAATTTGTGGAGGCAAG 287
DB 181 AATTACCGCACTGAAAAATGGGCTGTCTCTACTTATTTATTTTGTGCAATTTGTGGAGGCAAG 240
QY 288 AAATCAATATTGGAAGCTCTTATTTGAAAGGCTCCGCCCATCTCGACTGCAAGAAT 347
DB 241 AAATCAATATTGGAAGCTCTTATTTGAAAGGCTCCGCCCATCTCGACTGCAAGAAT 300
QY 348 GGATTTACAGCCTTGCAATTTAGCAGTTTACAGGATAATGCAGAAATGATCACTTCTCTG 407
DB 301 GGATTTACAGCCTTGCAATTTAGCAGTTTACAGGATAATGCAGAAATGATCACTTCTCTG 360
QY 408 CTTTCAGTGGAGCTGATATACAGAGTTGGATAGCGTGGCTCACTGCCCTCCATATT 467
DB 361 CTTTCAGTGGAGCTGATATACAGAGTTGGATAGCGTGGCTCACTGCCCTCCATATT 420
QY 468 GCTACAATAGCTGGCCACCTTAGAGGCTGTGTGTCTTGTGCAACATGGAGCTAATGTC 527
DB 421 GCTACAATAGCTGGCCACCTTAGAGGCTGTGTGTCTTGTGCAACATGGAGCTAATGTC 480
QY 528 AATATTCAAGATGAGTTTTTTTCACTCCATTCATATTGAGAGCTACTATGGACATGAA 587
DB 481 AATATTCAAGATGAGTTTTTTTCACTCCATTCATATTGAGAGCTACTATGGACATGAA 540
QY 588 CAGTAACTCGCTCTTTTGAATTTGTGTGTATGTAATGTAAGTGGTGAAGTTGA 647
DB 541 CAGTAACTCGCTCTTTTGAATTTGTGTGTATGTAATGTAAGTGGTGAAGTTGA 600
QY 648 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTTGAATATTGCAAAATCTTTGATG 707
DB 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTTGAATATTGCAAAATCTTTGATG 660
QY 708 GAAGAGGCGAGCAAGCAGATGTAATGCTCAAGATAATGGAAGCAATGCCACTCCAT 767
DB 661 GAAGAGGCGAGCAAGCAGATGTAATGCTCAAGATAATGGAAGCAATGCCACTCCAT 720
QY 768 TTCTGTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTCAAAAGTGAATTTGAA 827
DB 721 TTCTGTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTCAAAAGTGAATTTGAA 780
QY 828 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACTGCGATGCTACAT 887
DB 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACTGCGATGCTACAT 840
QY 888 GGCAGAAATTTCAAGTTGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947
DB 841 GGCAGAAATTTCAAGTTGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
QY 948 GAAAGACATCTTCAGTGAACACAGCTTTTTCATAGTGTGTTGACCTATGGCAAGACATGAC 1007
DB 901 GAAAGACATCTTCAGTGAACACAGCTTTTTCATAGTGTGTTGACCTATGGCAAGACATGAC 960
QY 1008 CTAGTCAAAATTTCTTCTGTATCAGATGTCATAAATCAACACCAAGGAGGATGG 1067
DB 961 CTAGTCAAAATTTCTTCTGTATCAGAAATGTCATAAATCAACCAAGGAGGATGG 1020
QY 1068 CACACTGGATTACACTCTGCTTGTCTACCGGTCACATTCGGCTGGTTCAGTTCTTACTG 1127
DB 1021 CACACTGGATTACACTCTGCTTGTCTACCGGTCACATTCGGCTGGTTCAGTTCTTACTG 1080

QY 1128 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGAGGCTCTAGTGGTGA AAA 1187
Db 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGAGGCTCTAGTGGTGA AAA 1140
QY 1188 GATGAGCAGACATGTTTGTATGTGGCTTATGAAAAGGGCATGATGCCATTGTCA CACTC 1247
Db 1141 GATGAGCAGACATGTTTGTATGTGGCTTATGAAAAGGGCATGATGCCATTGTCA CACTC 1200
QY 1248 CTGAAGCATTATAAGAGACCAAGATGAATGCGCTGTATGAATATCTCAGCCTGGA 1307
Db 1201 CTGAAGCATTATAAGAGACCAAGATGAATGCGCTGTATGAATATCTCAGCCTGGA 1260
QY 1308 GGAGATGGCTCCTATGTGTCTGTCCATCAGCTTGGGGAAGATTAAAGCATGACAAA 1367
Db 1261 GGAGATGGCTCCTATGTGTCTGTCCATCAGCTTGGGGAAGATTAAAGCATGACAAA 1320
QY 1368 GAGAGGCGAGATATCTCCTCTTAAGAGCTGGATTGCTTCAATTTCCATCTCAGCTC 1427
Db 1321 GAGAGGCGAGATATCTCCTCTTAAGAGCTGGATTGCTTCAATTTCCATCTCAGCTC 1380
QY 1428 TCAGAAATTTGAGTTCATGAGATATTGGCTCAGTCTCTTTGGGAAAGTATATAAGGA 1487
Db 1381 TCAGAAATTTGAGTTCATGAGATATTGGCTCAGTCTCTTTGGGAAAGTATATAAGGA 1440
QY 1488 CGATCAGAAATPAAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCAAG 1547
Db 1441 CGATCAGAAATPAAATAGTGGCTATAAAGCTTATCGAGCCAAATACCTACTGCTCAAG 1500
QY 1548 TCAGATGTCGATATGTTTGGGAGAGTGTCATCTCTGCCAGCTCAATCATCCCTGC 1607
Db 1501 TCAGATGTCGATATGTTTGGGAGAGTGTCATCTCTGCCAGCTCAATCATCCCTGC 1560
QY 1608 GTAATTCAGTTTGTGGTGTCTGTTGAATGATCCAGCCAGTTTGGCAATGTCACTCAA 1667
Db 1561 GTAATTCAGTTTGTGGTGTCTGTTGAATGATCCAGCCAGTTTGGCAATGTCACTCAA 1620
QY 1668 TACATATCAGGGGGTCTCTGTTCTCCCTCTTATGAGCAGAGAGGATCTTGATTG 1727
Db 1621 TACATATCAGGGGGTCTCTGTTCTCCCTCTTATGAGCAGAGAGGATCTTGATTG 1680
QY 1728 CAGTCTAAATTAATTTGTCAGTAGATGTTGCCAAGGCATGGAGTACCTTCAACCTG 1787
Db 1681 CAGTCTAAATTAATTTGTCAGTAGATGTTGCCAAGGCATGGAGTACCTTCAACCTG 1740
QY 1788 ACACAGCCAAATPATATCGTGAATGAAAGTCAATATCTTCTCTATGAGGATGGG 1847
Db 1741 ACACAGCCAAATPATATCGTGAATGAAAGTCAATATCTTCTCTATGAGGATGGG 1800
QY 1848 CATGCTGTGGTGGCAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1907
Db 1801 CATGCTGTGGTGGCAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
QY 1908 ATGACAAAACCACTGGGAACCTCGTTGGATGGCTCTGAGGTGTTTCACGAGTGCAC 1967
Db 1861 ATGACAAAACCACTGGGAACCTCGTTGGATGGCTCTGAGGTGTTTCACGAGTGCAC 1920
QY 1968 CGGTACACCATCAAGCAGATGTTCTCAGCTATGCTGTGCTGTGGGAAATTCACCT 2027
Db 1921 CGGTACACCATCAAGCAGATGTTCTCAGCTATGCTGTGCTGTGGGAAATTCACCT 1980
QY 2028 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGGAGCAGACATGGTTACCCAC 2087
Db 1981 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGGAGCAGACATGGTTACCCAC 2040
QY 2089 CACATCAGACCTCCCATGGCTATTCATTCGAGCCGATATCATCTCTGCTGATACGA 2147
Db 2041 CACATCAGACCTCCCATGGCTATTCATTCGAGCCGATATCATCTCTGCTGATACGA 2100
QY 2148 GGGTGGAGCGCATGCTCTGAGGAGAGACCCCAATTTTCTGAAGTTGTTCATGAAGTAGAA 2207
Db 2101 GGGTGGAGCGCATGCTCTGAGGAGAGACCCCAATTTTCTGAAGTTGTTCATGAAGTAGAA 2160
QY 2208 GAGTGTCTGCAACATTTGAGCTGATGTTCTCTGATCAGATCAAGTAACGAGCTGGTCTCTC 2267

Db 2161 GAGTGTCTGCAACATTTGAGCTGATGTTCTCTCTGATCAAGTAACAGCAGTGGTCTCTC 2220
QY 2268 TCACCTTCTTCTTCTTCTGATTTGCTGTGTAACCGGGGAGGACCTGCGCGGAGTCAATGTG 2327
Db 2221 TCACCTTCTTCTTCTTCTGATTTGCTGTGTAACCGGGGAGGACCTGCGCGGAGTCAATGTG 2280
QY 2328 GCAGCATTTAAGAGTCTGTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2387
Db 2281 GCAGCATTTAAGAGTCTGTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2340
QY 2388 TTGTCCCAAGTGTGGACAATATTCCTCTCAAGGTCTGTCTTTGGAGGAGATGAAAAGA 2447
Db 2341 TTGTCCCAAGTGTGGACAATATTCCTCTCAAGGTCTGTCTTTGGAGGAGATGAAAAGA 2400
QY 2448 AGTCTTCAATACACACCCATTGACAATATGGCTATGATCCGATCCCATGAGCTCAATG 2507
Db 2401 AGTCTTCAATACACACCCATTGACAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460
QY 2508 CATTTTCATTTTGGCGAAATAGTAGCAGCTTTTGAGCAGCAGC 2552
Db 2461 CATTTTCATTTTGGCGAAATAGTAGCAGCTTTTGAGCAGCAGC 2505

RESULT 5

AAS80720

ID AAS80720 standard; cDNA; 2788 BP.

XX AC AAS80720;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #16524.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABE16533.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

XX PT responsible for genetic disorders or other traits and to assess

XX PT biodiversity.

XX PS Claim 1; SEQ ID NO 16524; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders

involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pat_sequences

Sequence 2788 BP; 807 A; 585 C; 626 G; 770 T; 0 U; 0 Other;

Query Match 81.3%; Score 2458.4; DB 5; Length 2788;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2470; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 85 CTGATGAATGGAAGAAAAGTCAGTGAATCATATGTTATCACAATAGAAAAGATTAGAAG 144
DB 317 CTGATGAATGGAAGAAAAGTCAGTGAATCATATGTTATCACAATAGAAAAGATTAGAAG 376
QY 145 ATGACCTGCAGATCAAGGAAAAGAACTGACAGAACTTAAGGAATATATTTGGCTCTGATG 204
DB 377 ATGACCTGCCGATCAAGGAAAAGAACTGACAGAACTTAAGGAATATATTTGGCTCTGATG 436
QY 205 AAGCCTTCAGTAAAGTCAATTTAAATPACCGACTGAAAATGGCGTGTCTCTACTTCATT 264
DB 437 AAGCCTTCAGTAAAGTCAATTTAAATPACCGACTGAAAATGGCGTGTCTCTACTTCATT 496
QY 265 TATGTCGATTTGGGAGGCAAGAAATCAATATTCGAACCTCTATGTTGAAA-GGGCTC 323
DB 497 TATGTCGATTTGGGAGGCAAGAAATCAATATTCGAACCTCTATGTTGAAAAGGGGCTC 556
QY 324 CGCCCATCTCGACTGACAAGAAATGGAATTTACAGCCTTGCAATTTAGCAGTTTACAAGGAT 383
DB 557 CGCCCATCTCGACTGACAAGAAATGGAATTTACAGCCTTGCAATTTAGCAGTTTACAAGGAT 616
QY 384 AATCAGAAATGATCATCTCTGCTTCACAGTGGAGCTGATATACAGAGGTTGGATAC 443
DB 617 AATCAGAAATGATCATCTCTGCTTCACAGTGGAGCTGATATACAGAGGTTGGATAC 676
QY 444 GGTGGCCTCACTGGCCCTCCATATTTGCTACAAATAGCTGGCCACCTAGAGCTCTGATG 503
DB 677 GGTGGCCTCACTGGCCCTCCATATTTGCTACAAATAGCTGGCCACCTAGAGCTCTGATG 736
QY 504 CTGTTGCAACATGAGGCTAATGTCATATTCAGATGCAAGTGTCTTTTCTCACTCCATTCGAT 563
DB 737 CTGTTGCAACATGAGGCTAATGTCATATTCAGATGCAAGTGTCTTTTCTCACTCCATTCGAT 796
QY 564 ATTGCAGGCTACTATGACATGACAGGTAACCTGCTCTCTTTTGGAAATTTGGTCTGAT 623
DB 797 ATTGCAGGCTACTATGACATGACAGGTAACCTGCTCTCTTTTGGAAATTTGGTCTGAT 856
QY 624 GTAAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAAGGATTC 683
DB 857 GTAAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAAGGATTC 916
QY 684 TTGAAATTCGAAAACCTCTTGTATGGAAGAGGAGGAGGAGGAGATGTAATGCTCAAGAT 743
DB 917 TTGAAATTCGAAAACCTCTTGTATGGAAGAGGAGGAGGAGGAGATGTAATGCTCAAGAT 976
QY 744 AATGAAGACATGTCACCTCCATTTCTGTTCTGATTTGGACACCATGATATAGTTAAG 803
DB 977 AATGAAGACATGTCACCTCCATTTCTGTTCTGATTTGGACACCATGATATAGTTAAG 1036
QY 804 TATCTGTCGAAAGTGAATTTGGAGTTCAACCTCATGTTGTTAATATCATTCGAGATACC 863
DB 1037 TATCTGTCGAAAGTGAATTTGGAGTTCAACCTCATGTTGTTAATATCATTCGAGATACC 1096
QY 864 CCCTTACCTGCGATGCTACATGCAATTTGAGTTGGCAAGGAATCATCCAAATA 923
DB 1097 CCCTTACCTGCGATGCTACATGCAATTTGAGTTGGCAAGGAATCATCCAAATA 1156

QY 924 TCAGGAACAGAAAGTCTGACTAAGGAAAACATCTTCAGTGAAACAGCTTTTTCATAGTGT 983
DB 1157 TCAGGAACAGAAAGTCTGACTAAGGAAAACATCTTCAGTGAAACAGCTTTTTCATAGTGT 1216
QY 984 TGTAACCTATGGCAGAGCAATGACCTAGTCAAAATTTCTTCTTGATCAGAAATGCTATAAC 1043
DB 1217 TGTAACCTATGGCAGAGCAATGACCTAGTCAAAATTTCTTCTTGATCAGAAATGCTATAAC 1276
QY 1044 ATCAACACCAAGGAGGATGGGACACCTGGATACCTCTGCTTGTACACCGGTAC 1103
DB 1277 ATCAACACCAAGGAGGATGGGACACCTGGATACCTCTGCTTGTACACCGGTAC 1336
QY 1104 ATTGCGCTGTTTCAAGTTCTTACTGGATTAATGAGCTGATATGAATCTAGTGGCTTGTGAT 1163
DB 1337 ATTGCGCTGTTTCAAGTTCTTACTGGATTAATGAGCTGATATGAATCTAGTGGCTTGTGAT 1396
QY 1164 CCCAGCAGGCTCTAGTGGTGAAGAAAGATGACGACAGATGTTTGATGTGGCTTATGAAGAA 1223
DB 1397 CCCAGCAGGCTCTAGTGGTGAAGAAAGATGACGACAGATGTTTGATGTGGCTTATGAAGAA 1456
QY 1224 GGGCATGATGCCATTTGTCACTCTCTGAAGCACTTAAGAGACCAACAGATGAATGGCC 1283
DB 1457 GGGCATGATGCCATTTGTCACTCTCTGAAGCACTTAAGAGACCAACAGATGAATGGCC 1516
QY 1284 TGTAATGAATATTTCTCAGGCTGGAGGATGGCTCTATGTGTCTGTTCATCACCCCTTG 1343
DB 1517 TGTAATGAATATTTCTCAGGCTGGAGGATGGCTCTATGTGTCTGTTCATCACCCCTTG 1576
QY 1344 GGGAGGATTAAGAGCATGACAAAGAGAGGAGGAGATTTCTCTCTTAAGAGCTGATG 1403
DB 1577 GGGAGGATTAAGAGCATGACAAAGAGAGGAGGAGATTTCTCTCTTAAGAGCTGATG 1636
QY 1404 CTTTCAATTTCCATCTTCAGCTCTCAGAAATTTGAGTTCCATGAGATTTATGGCTCAGGT 1463
DB 1637 CTTTCAATTTCCATCTTCAGCTCTCAGAAATTTGAGTTCCATGAGATTTATGGCTCAGGT 1696
QY 1464 TCCTTTGGGAAAGTATATAAGACCATGACAGAAATAAATAGTGGCTATATAACGTTAT 1523
DB 1697 TCCTTTGGGAAAGTATATAAGACCATGACAGAAATAAATAGTGGCTATATAACGTTAT 1756
QY 1524 CGAGCCAATAACCTACTCTCCAAAGTCAGATGTTGGATATGTTTGGCGAGAGGTGTCATT 1583
DB 1757 CGAGCCAATAACCTACTCTCCAAAGTCAGATGTTGGATATGTTTGGCGAGAGGTGTCATT 1816
QY 1584 CTCTGCCAGCTCAATCATCCCTGCGTAATCAGTTTGTGGGTGCTTGTGTAATGATCC 1643
DB 1817 CTCTGCCAGCTCAATCATCCCTGCGTAATCAGTTTGTGGGTGCTTGTGTAATGATCC 1876
QY 1644 AGCCAGTTTCCCATTTGTCATCAATACATATCAGGGGGTTCTCTGTTCCCTCCCTCAT 1703
DB 1877 AGCCAGTTTCCCATTTGTCATCAATACATATCAGGGGGTTCTCTGTTCCCTCCCTCAT 1936
QY 1704 GAGCAGAAGAGGATCTTGATTTGCACTCTAAATTAATTTATTCAGATAGATGTTGCCAAA 1763
DB 1937 GAGCAGAAGAGGATCTTGATTTGCACTCTAAATTAATTTATTCAGATAGATGTTGCCAAA 1996
QY 1764 GGCATGGGTACCTTCACACCTGACACAGCCAAATATATACATCGCTTGACAGTGCAC 1823
DB 1997 GGCATGGGTACCTTCACACCTGACACAGCCAAATATATACATCGCTTGACAGTGCAC 2056
QY 1824 AATATTTCTTCTCTATGAGGATGGGCAATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTT 1883
DB 2057 AATATTTCTTCTCTATGAGGATGGGCAATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTT 2116
QY 1884 CTACAGTCTCTGATGAAGACAAACATGACAAAACAACTCGGGAACTCTCGCTTGGATGGCT 1943
DB 2117 CTACAGTCTCTGATGAAGACAAACATGACAAAACAACTCGGGAACTCTCGCTTGGATGGCT 2176
QY 1944 CTTGAGGTGTTACGAGTGCATCGGTACACCATCAAGAGAGATGTTCTTCAGCTATGCT 2003
DB 2177 CTTGAGGTGTTACGAGTGCATCGGTACACCATCAAGAGAGATGTTCTTCAGCTATGCT 2236
QY 2004 CTGTGTTCTGTGGAAATTTCTCACTGGCGAAATTTCCATTTGCTCATCTCAAGCCAGCGGT 2063

2237	Db		CTGTGCTGTGGGAAAATTCTCACTGGCGAAATTCATTGCGCTCATCTCAAGCCGACGGCT	2296
2064	Qy		GGCGGACGAGACATGGCTTACCACACATCAGACCTCCCATTTGGCTATTTCATTTCCCAAG	2123
2297	Db		GGCGGACGAGACATGGCTTACCACACATCAGACCTCCCATTTGGCTATTTCATTTCCCAAG	2356
2124	Qy		CCCATATCATCTCTGCTGATACGAGGGTGGAAACGATGTCCTGAAAGGAAGACCCGAATTT	2183
2357	Db		CCCATATCATCTCTGCTGATACGAGGGTGGAAACGATGTCCTGAAAGGAAGACCCGAATTT	2416
2184	Qy		TCCTGAAGTTGTCATGAGTTAGAGAGTGCTCTGCAACATTTGAGCTGATGTCCTCTGCA	2243
2417	Db		TCCTGAAGTTGTCATGAGTTAGAGAGTGCTCTGCAACATTTGAGCTGATGTCCTCTGCA	2476
2244	Qy		TCAAGTAAACAGCAGTGGGTCTCTCTCACACCTCTTTCTTTCTGATGCTGGTGAACCGG	2303
2477	Db		TCAAGTAAACAGCAGTGGGTCTCTCTCACACCTCTTTCTTTCTGATGCTGGTGAACCGG	2536
2304	Qy		GGAGACCTGGCCGGAGTCATGGGAGCATTAAGAGTCGTTTTGCAATTTGGAATATGCT	2363
2537	Db		GGAGACCTGGCCGGAGTCATGGGAGCATTAAGAGTCGTTTTGCAATTTGGAATATGCT	2596
2364	Qy		CTAAATGCAAGGTCTATGTCTGCTTTGTCGCCAAAGTCTGGGCAATATTCCTCTCAAGGT	2423
2597	Db		CTAAATGCAAGGTCTATGTCTGCTTTGTCGCCAAAGTCTGGGCAATATTCCTCTCAAGGT	2656
2424	Qy		CTGTCTTTGGAGGAGATGAAGAAGTCTTCAATACACACCATTTGCAAAATATGGCTAT	2483
2657	Db		CTGTCTTTGGAGGAGATGAAGAAGTCTTCAATACACACCATTTGCAAAATATGGCTAT	2716
2484	Qy		GTATCCGATCCCATGAGCTCAATCCATTTTTCATTCTTGGCCGAAATAGTAGAGAGCTTTTGAG	2543
2717	Db		GTATCCGATCCCATGAGCTCAATCCATTTTTCATTCTTGGCCGAAATAGTAGAGAGCTTTTGAG	2776
2544	Qy		GACAGCAGCTGA	2555
2777	Db		GACAGCAGCTGA	2788

RESULT 6
AAA47608
ID AAA47608 standard; cDNA; 3026 BP.
XX
XX AC AAA47608;
XX
XX
XX 20-OCT-2000 (first entry)
XX
XX DE Rat CARK (Cardiac related Ankyrin-Repeat Protein Kinase) cDNA.
XX
XX Cardiac related ankyrin repeat protein kinase; CARK; cytoskeleton;
KW cardiac cell growth factor receptor; cell differentiation; modulator;
KW regulator; detection; cellular proliferation; cardiovascular disorder;
KW heart failure; hypertension; cancer; sarcoma; ds.
KW

Pt	Raju J;
XX	WPI; 2000-431275/37.
DR	P-PSDB; AAB01474.
XX	
XX	New polynucleotide encoding cardiac-related ankryrin-repeat protein kinase, useful for treating disorders such as cardiovascular disorders,
PT	e.g. heart failure and cell differentiation disorders, e.g. cancer.
PT	
XX	Claim 1; Fig 5; 16pp; English.
PS	
XX	CARK polypeptides are regulators of signal transmission from cellular receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton function. They also act as modulators of cellular differentiation and cell death. Nucleic acids, or their fragments encoding CARK polypeptides are useful for detecting CARK nucleic acids especially mRNA, in a sample. CARK is useful for treating disorders associated with upregulation or downregulation of cellular proliferation such as, cardiovascular disorders (heart failure and hypertension) and disorders associated with cell differentiation such as cancer and sarcoma
CC	
XX	Sequence 3026 BP; 778 A; 751 C; 743 G; 754 T; 0 U; 0 Other;
SQ	
	Query Match 66.4%; Score 2010; DB 3; Length 3026;
	Best Local Similarity 80.6%; Pred. No. 0;
	Matches 2425; Conservative 0; Mismatches 570; Indels 13; Gaps 6;
QY	16 CGGCCCTCGAAGAGGAAACTTATATTAATGGGAATTATAAATCTAGACCAACC 75
Db	29 CAGCACGAGAAGCAAAGACTCTTTAAATGGGAATTACAATTCAGACCAACAC 88
QY	76 AAATCTTGACTCATCAATGGGAAGAAAAAGTCAGTGAATCATATGTATTATCAATAGAAA 135
Db	89 AGACTTGTCGTGATGAATGGAAGAAAGTAGTGAATCTTACGCCTATTATCATAGAAA 148
QY	136 GATTAGAAGATGACTCGAGATCAGGAAAAAGAACTGACAGAACTAGGAATATTTG 195
Db	149 GGCTGGAGGATPAACCTCGCATCAAGAAAAATGAATTTCAAGAACTAAGGCACATCTTTG 208
QY	196 GCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTCACGCATCAAAATGGGTGCTC 255
Db	209 GCTCTGATGAAGCCTTCAGTGAAGTCAATTTAAATTCACGCATCAAAATGGGTGCTC 268
QY	256 TACTTCAATTAATGTCATTTGTCAGGAGCAAGAAATCAATATTCGAATCTTATGTTGA 315
Db	269 TGCTACACTCTGCTGTCTGTGGCGGCAACAAGTCACATATCGCTGCCCTTATGTTAA 328
QY	316 AAGGCTCGGCCATCTCGATCGACAGAAATGGATTTACAGCCTTGCATTTAGCAGTTT 375
Db	329 AAGGCTCGGCCATCGAGACTGACAGAAATGGATTTCCAGCTCTGCACCTGGCCGTTT 388
QY	376 ACAAGGATAATCGCAATTCGATCACTTCTCGCTTCAAGTGGAGCTGATATACAGCAGG 435
Db	389 ACAAGGACAGCCCGAACTTATCACTTCACTGTTGACAGCGAGCAGATGTCAGCAAG 448
QY	436 TTGGATAGGTGGCTCACTGCCCCATATTCCTAATAGCTGCGCCACCTAGAGGCTG 495
Db	449 TGGGATAGGTGGCTCACTGAGCCCTCCACATAGCTGCAATAGCTGGACCCAGAGGCTG 508
QY	496 CTGATGCTGCTTGGCAATGAGAGCTAATGTCAAATTCAGATGCGAGTTTTTTTTCATC 555
Db	509 CAGAAGTGCCTGCTACACATGGGGCCATGTCGATGTTCAAGATCCGCTCTTCTCACCC 568
QY	556 CATTCGATATTGACGGCTACTATGGACATGAACAGGTAACTCGCCTTCTTTTGAATTTG 615
Db	569 CACTGCATATTGACGCTCTACTATGGGCACGACAGGTAAACAGTGTCTCTTTTGAAGTTG 628
QY	616 GTGCTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAA 675
Db	629 GTGCTGATGTCATGTAAGCGGTGAAGTTGGGACAGGCTCTGACCTTGGCCCTCTGCAA 688
QY	676 AAGGATCTTGAATATTGCAAAACTCTTGTATGGAAGAGGACGAAAGCAGATGTGAATG 735

Db 689 AGGGCTTCTTCAACATTTGTGAAGCTCTCTGCTAGAGNAGGAGCAAGCAGATGTGAACG 748
Qy 736 CTCAGATATGAAGACCATGTGCCACTCCATTTCTGTTCTCGATTTGGACACCATGATA 795
Db 749 CTCAGGCAATGAAGACCATGCTCCCTCTGCACTTCTGTTCTCGATTTGGACACCAATA 808
Qy 796 TAGTAAAGTATCTGCTGCAAGTGAATTTGAAGTTCAACCTCATGTGTTTAAATATCTATG 855
Db 809 TAGTGAGTACCTGCTCCAGAGTGACTTAGAGTCCAGCCTCAGTCATTAACAATCTATG 868
Qy 856 GAGATACCCCTTACACTCGCATGTACAAATGGCAAAATTTGAAGTTGGCAAGGAAATCA 915
Db 869 GTGACACTCTCTTGCACCTCGCATGTACAAATGGAAATTTGAAGTTGGCAAGGAAATG 928
Qy 916 TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTCACTGAGAAACAGCTTTTC 975
Db 929 TCCAGGTAAACAGGAACTGAAAGTCTGACTAAGGAAACATCTTCAAGGAGACAGCTTTTC 988
Qy 976 ATAGTGTCTGTACTATGGCAAGCATGACCTAGTCAAAATTTCTTTGATCAGAATG 1035
Db 989 ACAGTGTCTGTACTATGGCAAGCAATTCACCTGTGCAAAATTTCTTTGATCAGATG 1048
Qy 1036 TCATAAATCAACCAAGGAGGATGGGACACTGGATTAACACTCTGCTTGTCTAC 1095
Db 1049 CTGTGAACATTAACCAAGGAGGAGATGGGACAGAGTTGCACTCTGCTTGTCTACC 1108
Qy 1096 ACAGTCACTTGGCTGTGTTCAAGTCTTCTTCTGATTAATGGAGTGAATGAATCTAGTGG 1155
Db 1109 ACAGCATATCCGCTGTGTTCAAGTCTTCTTCTGATTAATGGTGCAGATATGAATCTGTG 1168
Qy 1156 CTGTGTATCCAGAGAGTCTAGTGTGGTGAAGAAAGATGAGCAGACATGTTTGTGTGGGCT 1215
Db 1169 CTGTGTATCCAGCAGTCTAGTGTGGTGAAGAAAGATGAGCAGACATGTTTGTGTGGGCT 1228
Qy 1216 ATGAAAAGGCGATGATGCAATTTGTACACTCTGGAAGCATTAAGAGACCAAGATG 1275
Db 1229 ACGAAGAAAGACATGATGCCAATTTGTACCTCTGAAAGCATTAAGAGACCCAGAGG 1288
Qy 1276 AATTGCCCTGTAAATGAATTTCTCAGCTGGAGAGATGGCTCTATGTGTCTGTTCCT 1335
Db 1289 AGCTGCCATGTAAACGAATATCCAGCCTGGAGAGATGGCTCTATGTGTCTGTTCCT 1348
Qy 1336 CACCTTTGGGAGATTAAGCATGACAAAGAGAGGAGCATATCTCTCCCTCAAGAG 1395
Db 1349 CCCCCTTTGGGCAAGATTAAGCATGACAAAGAGAGGAGATGTTCTCTCCCTGAGG 1408
Qy 1396 CTGGAATTCCTTCAACATTTCCATCTTCAAGCTCTCAGAAATTTGAGTTCCATGAGATTTG 1455
Db 1409 CTGAACCTACCTCCGCTTCCATCTCCAACTCTCCGAATTCGAGTTCCACGAGATATCG 1468
Qy 1456 GCTCAGGTTCTTTTGGGAAAGTATATAAGGAGCGATGCGAATAAATAATAGTGGCTATA 1515
Db 1469 GCTCGGTTCTTTTGGGAAAGTCTATAAAGGGCGATGCGAATAAATAATAGTGGCGATCA 1528
Qy 1516 AACGTTATCGAGCAATACCTACTGTCTCAAGTCAGATGAGATGTTTTCGCCAGAGG 1575
Db 1529 AACGATACCGAGCCACACCTACTGCTCCAAAGTCAGACGTGAGATGTTTTCGCCAGAGG 1588
Qy 1576 TGTCCATCTCTGCCAGCTCAATCATCCCTCGTAAATTCAGTTTGGGCTGTCTCTTGA 1635
Db 1589 TGTCCATCTCTGCCAGCTCAACCAACCCCTCGTGGTTCAGTTTGGGCTGTCTCTCTG 1648
Qy 1636 ATGATCCAGCGAGTTTGGCAATTTGTCATCTCAATATACATATCAGGGGTTCTCTGTTCTCC 1695
Db 1649 ATGACCCAGCTAGTTTGGCAATTTGTCATCTCAGTACATTTTCAAGAGGCTCCCTGTTCTCC 1708
Qy 1696 TCCCTTCAAGAGAGAGGATTTCTTGAATTTGAGTCTAAATTAATTAATTCAGTAGATG 1755
Db 1709 TGTCTCATGAACAGAGAGATTTCTTGAATTTGAGTCTAAATTAATTAATTCAGTAGAGG 1768
Qy 1756 TTGCAAGGAGTGTGAGTACCTTCAACACTGACACAGCCATTAATATCATCTGATCTGA 1815
Db 1769 TTGCCAAGGCGATGGAGTACCTCTGACAGCTTTGACCCAGCCCAATCATACACCGGACCTGA 1828

Qy 1816 ACAGTCACAATATTTCTTCTATGAGGATGGGATGCTGTGTGGCAGATTTTGGAGAT 1875
Db 1829 ACAGCCACAATATTTCTGCTCTATGAGGATGGCCATGCTGTGTGGCAGATTTTGGAGAT 1888
Qy 1876 CAAGATTTCTACAGTCTCTGGATGAAGACAAATGACAAAAAACAACCTGGGAACTCCGCT 1935
Db 1889 CAAGATTTCTGCACTCTCTGGATGAAGACAAATGACAAAAAACAACCTGGGAACTCCGCT 1948
Qy 1936 GGATGGCTCTCAGGTTGTTTCAAGCAGTGCATCTGGTACACCATCAAAAGCAGATGTTTCA 1995
Db 1949 GGATGGCCCTCAGGTTGTTTCAACAGTGCACGAGATACACCATCAAGGCTGATGTTTCA 2008
Qy 1996 GCTATGCTCTGCTGTGGGAAATTTCTCACTGGGAAATTTCCATTCGCTCATCTCAAGC 2055
Db 2009 GTTACTCCCTGCTGTGGAGCTCTCACTGGAGAAATTTCCATTCGCTCATCTCAAGC 2068
Qy 2056 CAGGGCTGGGAGCAGACATGGCTTTACCAACATCAGACCTCCCATTTGGCTATTTCA 2115
Db 2069 CAGCGCTGCAGCAGACAGATATGGCGTATCACACATCAGACCGCCCATCGGCTATTCCA 2128
Qy 2116 TTCCCAAGCCCATATCATCTCTGATAGAGGGTGGAAACGATGCTCTGAGGAGAC 2175
Db 2129 TCCCAAGCCCATCTCATCTCTGATAGGGCTGGAAATGATGCTCTGAGGAGACGAC 2188
Qy 2176 CCGAATTTTCTGAAGTTGTCATGAAGTTAGAAGTGTCTCTGCAACATTTGAGCTGATGT 2235
Db 2189 CAGAGTTCTCTGAAGTCGTTAGCAAACTGGAGAGTGCCTATGCAATGTGGAGCTCATGT 2248
Qy 2236 CTCTGATCAAGTAAACAGAGTGGTCTCTCTCACTTCTTCTTCTTCTGATTTGGCTGG 2295
Db 2249 CTCAGCATCAAGTAAACAGAGTGGTCTCTCTGCACTTCTTCTTCTTCCGATTCCTG 2308
Qy 2296 TGAACCGGGAGACCTGGCCGAGTCATGTGGCAGCATTAAGAAGTCGTTTCGAATTTGG 2355
Db 2309 TGAGCCGGGAGAGGCTTGGCCGGAGCCAAGTGGCAGCTTACGGAGCCGTTTGTGATGG 2368
Qy 2356 AATATGCTCTAAATGCAAGTCTCTATGCTGTTTGTCCAAAGTGTGACAAATATTCCT 2415
Db 2369 AGTATGCCCTAAATGCAAGTCTCTATGCTGGGTGCTCCAAAGTGTGGAAACACACTCTA 2428
Qy 2416 CTCAAGTCTGCTTTGGAGGAGATGAAGAAGTCTTCAATACACACCATTTGACAAAT 2475
Db 2429 ATCCGGCCCTGCTTTTGGAGGAGATGAATAGGAGCACCAGATTTCACTGTTGACAAAT 2488
Qy 2476 ATGGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCAATTTTCCGAAATAGTAGCA 2535
Db 2489 ACGCTATGTGCTGATCCCATGAGCTGACGACCTTCACTCCCGCAAGACGACAGCA 2548
Qy 2536 GCTTTGAGGACAGCTGACAGCATTCGGCGTATACCTAAGAGAGTGTTCCTCCCGNA 2595
Db 2549 ACTTTGAGGACAGCAACTGACAG-GTCTGGCATACACCTAAGGGGCTCTCCCATCAGG 2607
Qy 2596 CTGACAGCAAGATTCACACCGCAAGCTGGCTTCCAACTATATAACATTTTACTCTCAA 2655
Db 2608 CTGACAGCATGATTTTACCATGGCAGGCTTGTCTTCCAAATTAACGGCCCTGCCCTCTG 2667
Qy 2656 AGGTCTCTTAAATTTGGGCTTGTGTTTATCTGTCTTATTTAAATTTCCCACTAATAGCAGG 2715
Db 2668 AGGT-TTCTTCAAAATCGTCTGCTTATTTAAAGTCTGTTTAAATTTCCCTTCTACAGGACAG 2726
Qy 2716 CTTTGGATTTTGCCTAAGGAATAATGCAAAAGAACCAAGACAGAAATGATATATGAAGA 2775
Db 2727 GCTTTGACTCATGCC--AAGCCTGAAGTGTCAAGAGCAGATACAGAAATGTCATGAGGA 2784
Qy 2776 ATTGTTTTTAAATTTGTAATTAATAAAAAAATTTAGATCGTTTACTTTGAAATTTGGAGCCTA 2835
Db 2785 ATTGTTCTTAGTTTGATATTTAAAGCCCTTAATTCCTGGGCTGGGGTTCAA---TC 2840
Qy 2836 AGTCTGTGGTGGACATATAATATATGTTTTCCTGGGCTGAATATGATAGCTTGTGTT 2895
Db 2841 TGTGTAGATAGTGGGTTGACCCCTATGTAATTTGTAGACCAAACTGTGTGGGCTTGTGTT 2900

Db 749 CTCAGGCAATGAGACCAGTCCCTCTGTCACCTTCTGTTCTCGATTGGACACCAATA 808
Qy 796 TAGTTAAGTATCTGCTGCCAAAGTGATTTGGAAGTTCAACTCATGTTGTTAATATCTATG 855
Db 809 TAGTGAGCTACCTGCTCCAGAGTGACTTAGAGTCCAGGCTCAGCTCATTAACATCTATG 868
Qy 856 GAGATACCCCTTACACCTCGGCATGTACAAATGGCAAAATTTGAAGTTGGCAAGGAAATCA 915
Db 869 GTGACATCTCTTTGCACTCGCATGCTCAATGGAAATTTTGAAGTTGGCAAGGAAATG 928
Qy 916 TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTTCAGTGAAAACAGCTTTTC 975
Db 929 TCCAGGTAACAGGAACTGAAAGTCTGACTAAGGAAACATCTTCAGCGAGACAGCTTTTC 988
Qy 976 ATAGTGCTTACTTACCTGAGAGCAATTCAGCTAGTTCACCTAGTTCAAAATTTCTTCTTGATCAGAATG 1035
Db 989 ACAGTGCTTGATCTATGGAAGAAACATTGACCTGGTCAAAATTTCTTCTTGATCAGAATG 1048
Qy 1036 TCATAAACATCAACCAACCAAGGAGGATGGGCACACTGGATATACACTCTGCTGCTTACC 1095
Db 1049 CTGTGAACATTAACCAACCGAGGAAGAGATGGGCACACAGGATTCACACTCTGCTTGTACC 1108
Qy 1096 ACAGTACATTCGGCTGGTTTCAAGTCTTACTGATTAATGGAGCTGATATGAATCTAGTGG 1155
Db 1109 ACGCCCATATCCGCTGGTTTCAAGTCTTACTTGATTAATGGTGAGATATGAATCTTGTGG 1168
Qy 1156 CTTGTGATCCAGCAGGCTTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTATGTGGGCTT 1215
Db 1169 CTTGTGATCCAGCAGGCTTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTATGTGGGCTT 1228
Qy 1216 ATGMAAAGGCGATGATGCCATTTGCACACTCTCGAAGCATTAAGAGACCAAGATG 1275
Db 1229 ACAGAAAGGACATGATGCCATTTGTACACTCTGAAAGCACTACAGAGACCCAGGAGG 1288
Qy 1276 AATTGCCCTGTAATGATATTTCTCAGCTGGAGGAGATGGCTCCTATGTCGTTCAT 1335
Db 1289 AGCTGCCATGTACGAATATTTCCAGCTGGAGGAGATGGCTCCTATGTCGTTCCTT 1348
Qy 1336 CACCTTTGGGAGATTAAGCATGACAAAAGAGAGGACAGATATTCCTCCTTAAGAG 1395
Db 1349 CCCCCTTGGGCAAGATTAAGCATGACAAAAGAGAGGACAGATTTCTCCTCCTGAGG 1408
Qy 1396 CTGGATTGCCTTACATTTCCATTTCAAGTCTCAGAAATGAGTCCATGAGATATG 1455
Db 1409 CTGAATACCCCTCCCGCTTCCATCTCCAACTCTCCGAAATCAGATTCACAGAGATATCG 1468
Qy 1456 GCTCAGGTTCTTTTGGGAAAGTATATAAGGACGATGCAGAAATAAAATAGTGGGTATAA 1515
Db 1469 GCTCGGTTCTTTTGGGAAAGTCTATAAGGGCGATGCAGAAATAAAATAGTGGGATCA 1528
Qy 1516 AACGTTATCGAGCAATACCTACTGCTCCAGTCCAGATGAGATATGTTTGGCCAGAGG 1575
Db 1529 AACGATACCGAGCAACACCTACTGCTCCAAAGTCAGACGTGGATATGTTTGGCCAGAGG 1588
Qy 1576 TGTCCATTTCTGCGAGCTCAATCATCCCTCGCTAATTCAGTTTGGGCTGCTGCTTGA 1635
Db 1589 TGTCCATTTCTGCGAGCTCAACCAACCCCTGCTGCTTCAAGTTTGGGCTGCTGCTG 1648
Qy 1636 ATGATCCCAACGCAATTTGCCATTTGTCATCAATACATATCAGGGGTTCTCTGTTCTCCC 1695
Db 1649 ATGACCCCACTGATTTGCCATTTGTCATCACTCAGTACATTTTTCAGGAGGCTCCCTGTTCTCCC 1708
Qy 1696 TCCTTCATGACGACAGAGAGATTTGATTTGAGTCTAAATTAATTAATGTCAGTAGATG 1755
Db 1709 TGCTTCATGAACAGAGAGAAATTTCTTGACTTTGAGTCTAAATTAATTAATGTCGGTAGAG 1768
Qy 1756 TTGCCAAAGGCATGGAGTACCTTCAACACCTGCACAGCCAAATTAATACATCGTGACTTGA 1815
Db 1769 TTGCCAAAGGCATGGAGTACCTGCACAGCTTGACCCAGCCCAATACACACCGACCTGA 1828
Qy 1816 ACAGTACAAATATCTTCTATGAGATGGGATGCTGCTGGTGGCAGATTTTGAGAAAT 1875
Db 1829 ACAGCCCAATATTTCTGCTCTATGAGGATGGCCATGCTGCTGGCAGATTTTGAGAAAT 1888

Qy 1876 CAAGATTTCTACAGTCTCTGGATGAAGACAAACATGACAAAACAACTGGGAACCTCGGT 1935
Db 1889 CAAGATTTCTGCAAGTCCCTGGATGAAGACAAACATGACAAAGCAGCAGGAAACCTGCGCT 1948
Qy 1936 GGATGGCTCTGAGAGTGTTCACGCAAGTGCACCTCGGTACACCACTCAAGAGCAGATGCTTCA 1995
Db 1949 GGATGGCCCTGAGGTGTTTACACAGTGCACGAGATACACCATCAAGGCTGATGCTTCA 2008
Qy 1996 GCTATGCTCTGCTGTCTGTTGGGAAATTTCTCACTGGGGAATTTCCATTCGCTCATCTCAAGC 2055
Db 2009 GTTACTCCCTGTCTGTTGGGAGCTCTCACTGGGAAATTTCCATTCGCTCATCTCAAGC 2068
Qy 2056 CAGCGCTGGCAGCAGACATGGCTTACCACACATCAGACCTCCCATTTGGCTATTTCCA 2115
Db 2069 CAGCGCTGCAGCAGCAGATATGGGCTATACACATCAGACCGGCCCATCGGCTATTTCCA 2128
Qy 2116 TTCCAAAGCCCATPATCATCTCTGCTGATACGAGGTGGAAACGCACTGCTCTGAAGGAAGAC 2175
Db 2129 TCCCAAGCCCATCTCATCCCTGCTGATACGGGGTGGAAATGCATGCTCTGAAGGACGAC 2188
Qy 2176 CCGAATTTTCTGAAGTGTCTGAAGTTAGAGAGTGTCTCTGCAACATTCCTGAGCTGATGT 2235
Db 2189 CAGAGTTCTCTGAAGTCTGTTAGCAAACTGGAGAGTGCCTATGCAATGTGGAGCTCATGT 2248
Qy 2236 CTCTGCATCAAGTAACAGCAGTGGGTCTCTCTCACCTTTCTTTCTTCTGATTCCTGG 2295
Db 2249 CTCAGCATCAAGTAACAGCAGTGGGTCTCTGTCACCTTCTCTTCTTCCGATTCGCTGC 2308
Qy 2296 TGAACGGGGAGACCTGGCCGGAGTCATGTGGCAGCATTAAGAAAGTGGTTTCGAATGG 2355
Db 2309 TGAGCCGGGAGGGCTTGGCCGGAGCACGTGGCAGCTTACGGAGCCGCTTTTGATTTGG 2368
Qy 2356 AATATGCTCTAAATGCAAGTCTCTATGCTGTCTTGTCCCAAAGTCTGGCACAATATTCCT 2415
Db 2369 AGTATGCCCTAAATGCAAGTCTCTATGCTGGGTGGTCCCAAGTGTGGAAACACACTCA 2428
Qy 2416 CTCAGGTCTGTCTTTGGAGGAGATGAAGAAAGTCTTCAATACACACCCATTTGACAAAT 2475
Db 2429 ATCCGGGCTGTCTTTGGAGGAGATGAATAGGAGCACCCAGTATTTCAACTGTTGACAAAT 2488
Qy 2476 ATGGCTATGATCCGATCCCATGAGCTCAATGCATTTTCATTTTCCGGAATATGATCA 2535
Db 2489 ACGCTATGTCGTCTGATCCCATGAGCTGACGACCTTCACCTCCGCCCAAGACAGACGA 2548
Qy 2536 GCTTTGAGGACAGCAGCTGACAGCATTCGGCGTATACCTTAAGGAGAGTTTTTCGCCGAA 2595
Db 2549 ACTTTGAGGACAGCAACTGACAG-GTCTGGCATACACCTAAGGGGCGTCTCCCCATCAGG 2607
Qy 2596 CTGACAGCAAGATTCACACCGGCAAGCTGGCTTCCAACTATAACATTTTACTCTCA 2655
Db 2608 CTGACAGCATGATTTTACCCATGGCAGGCTTGTCTCCAAATTAACGCCCTCGCCCTCTG 2667
Qy 2656 AGGTCTCCTTAAATTTGGGCTTGTTTTACTTGTCTCTTAAATTTCCCACTATTAGCAGG 2715
Db 2668 AGGT-TTCITTCAAATCGTCTTGCTTATTTCTAGCTCGTTTATTTCCCTTCTACAGACAG 2726
Qy 2716 CTTTGGATTTGGCTTAAGGAATATATGCAAAAGAACCAAGACAGATGTATATGAAGA 2775
Db 2727 GCTTTGACTCATGCC--AAAGCCTGAAGTGTCAAGAGCAGATACAGAATGTGCATGAGGA 2784
Qy 2776 ATTGTTTTTAATTTTCTAAATTAATAAAAAAATTTAGATCGTTTACTTGGAAATGGAGCCTA 2835
Db 2785 ATTGTTCTTAGTTTGATATTAAGCCCTTAATTCCTGGGCTGGGTTCAAA---TC 2840
Qy 2836 AGTCTGTGTGGACAGATAAATATATGTTTTTCTGGGCTGAATATATGATAGATTTGTGT 2895
Db 2841 TGTGTAGATAGCTGGGTTGACCCCTTATGATTTTGTAGACCAAACTGTGTGGGCTTGTGT 2900
Qy 2896 TGACAG-CTATGGGTTTATTTCTTAGACATTTGTCATTTTCTTCTCATATATCTACT 2954
Db 2901 TGAGGGTCTCCTTGGGTTTCTTAAAAAACAAGCTGGCTGATTTTATCTCTCTTCCCTTT 2960

QY	2955	TCTAGTGTCCACCTCTGTGATTAAAGATCTTTGGTGAATAGAAAAAAGAAAAA 3014	Db	61	AGTGAATCTTACGCTTATATATAGAAAAGGCTGGAGGATAACCTGCAGATCAAGAAAAAT 120
Db	2961	----GTTGTTACTTCTGTGATTAAGATCTCTTGGTGATCTAGAAAAAAGAAAAA 3016	QY	168	GAACTGACAGAACTAAGGATATATTGCTCTGTAGTGAAGCCCTTCAGTAAAGTCAATTTA 227
QY	3015	AGGGCGGC 3022	Db	121	GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCCTTCAGTGAAGTCAATTTA 180
Db	3017	GGGCGGCC 3024	QY	228	AAATTACCGCACTGAAAATGGGTGCTCTACTTTCATTTATATGTTGCAATTTGTGGAGGCAAG 287
RESULT 8			Db	181	AAATTACCGCACAGAGCGTGGCTGCTCCCTGCTACACCTCTGCTGTGTGTGTGTGGCGCAAC 240
AAA47609			QY	288	AAATCACATATTGAACTCTTATGTTGAAGGGCTCCGCCATCTTCGACTGCACAGAAAT 347
ID AAA47609 standard, cDNA; 2505 BP.			Db	241	AAGTCACATATCCGTGCCCTTATGTTAAAGAGGCTCCGTCCATCCAGACTGACGAGAAT 300
AC AAA47609;			QY	348	GGATTTACAGCCTTGCATTTTAGCAGTTTACAAAGGATAATGCAGAAATTGATCACTTCTCTG 407
XX 20-OCT-2000 (first entry)			Db	301	GGGTTTCAGCTCTGCACCTGGCCGTTTACAGGACAGCCCGGAACTTATCACTTCACTG 360
XX Rat CARX (Cardiac related Ankyrin-Repeat Protein Kinase) cDNA.			QY	408	CTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 467
XX Cardiac related ankyrin repeat protein kinase; CARX; cytoskeleton;			Db	361	TTGCACAGCGGAGCAGATGTTTCAAGTGGGATAGCGTGGCTCACAGCCCTCCACATA 420
KW cardiac cell growth factor receptor; cell differentiation; modulator;			QY	468	GCTACATAGCTGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGGAGCTAATGTC 527
KW regulator; defection; cellular proliferation; cardiovascular disorder;			Db	421	GCTCAATAGCTGGACACCCAGAGGCTGCAGAACTGCTACACATGGGCGCAATGTG 480
KW heart failure; hypertension; cancer; sarcoma; ds.			QY	528	AATATTCAGATGCACTTTTTCATCTGATTTGATGCGGTGCTATGCGAGTATCTATGGACATGAA 587
XX Rattus norvegicus.			Db	481	AATGTTCAAGATGCGGTCTTCTTCAACCCACTGCACATTTGCAGCCCTACTATATGGGACGAG 540
Key Location/Qualifiers			QY	588	CAGTAACTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTGGGA 647
1..2505			Db	541	CAGTAACTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTGGGA 600
/*tag= a			QY	648	GATAGACCTCCACCTAGCATCTGCAAAAGGATTTTGAATTTGCAAAATCTTTGATG 707
/product= "Rat CARX"			Db	601	GACAGGCTCTGCACCTGGGCTCTGCAAAAGGCTTCTTCAACATTTGTGAACCTCTGCTGA 660
WO200034330-A1.			QY	708	GAAGAAGGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGCTCCACTCCAT 767
15-JUN-2000.			Db	661	GAAGAAGGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGCTCCACTCCAT 720
10-DEC-1999; 99WO-US029465.			QY	768	TTCTGTTCTCGATTTGGACACCATGATATAGTTAACTATCTGCTGCAAAAGTGAATTTGCAA 827
11-DEC-1998; 98US-0111938P.			Db	721	TTCTGTTCTCGATTTGGACACCATGATATAGTTAACTATCTGCTGCAAAAGTGAATTTGCAA 780
14-APR-1999; 99US-00291839.			QY	828	GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGACCTGCATGCTACAT 887
XX (MILL-) MILLENNIUM PHARM INC.			Db	781	GTCCAGCCTCAACGCTAATCACTATGTTGACATCTTCTTGCACCTGCATGCTACAT 940
XX Raju J;			QY	888	GGCAAAATTCAGTTGCGCAAGGAAATCATCCAAATATCAGGACACAGAAAGTCTGACTAAG 947
XX WPI; 2000-431275/37.			Db	841	GGAAATTTGAAGTTGCGCAAGGAAATTCAGGTTAACAGGAACTGAAAGTCTGACTAAG 900
XX New polynucleotide encoding cardiac-related ankyrin-repeat protein			QY	948	GAAGAATCTTTCAGTGAACACAGCTTTTCAATGATGTTGTAACCTATGGCAAGAGCTTTGAC 1007
kinase, useful for treating disorders such as cardiovascular disorders,			Db	901	GAAGAATCTTTCAGTGAACACAGCTTTTCAATGATGTTGTAACCTATGGCAAGAGCTTTGAC 960
e.g. heart failure and cell differentiation disorders, e.g. cancer.			QY	1008	CTAGTCAAAATTTCTTTGATCAGATGTCATAAATCAACCAACCAAGGAGGATGGG 1067
Claim 1; Page 155-159, 161pp; English.			Db	961	CTGTGCAAAATTTCTTTGATCAGATGTCATAAATCAACCAACCAAGGAGGATGGG 1020
CARX polypeptides are regulators of signal transmission from cellular			QY	1068	CACACTGGAATTTACACTCTGCTTGTCTACCGGTGCACATTCGCTGGTTCAGTTCCTTACTG 1127
receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton			Db	1021	CACAGGATTTGCACTCTGCTTGTCTACCGGCTATATCCGCTGGTTCAGTTCCTTACTT 1080
function. They also act as modulators of cellular differentiation and			QY	1128	GATTAATGGAGTGTATATGATCTAGTGGCTTGTGATCCCGAGGCTAGTGGTGAAGAA 1187
cell death. Nucleic acids, or their fragments encoding CARX polypeptides			Db	1081	GATTAATGGAGTGTATATGATCTAGTGGCTTGTGATCCCGAGGCTAGTGGTGAAGAA 1140
are useful for detecting CARX nucleic acids especially mRNA, in a sample.			QY	1188	GATGAGCAGACATGTTTGTGCTTATGAAAAAGGCGCATGATGCCATTTGTCAACATCT 1247
CARX is useful for treating disorders associated with upregulation or					
downregulation of cellular proliferation such as, cardiovascular					
disorders (heart failure and hypertension) and disorders associated with					
cell differentiation such as cancer and sarcoma					
Sequence 2505 BP; 648 A; 640 C; 617 G; 600 T; 0 U; 0 Other;					
Query Match 62.6%; Score 1892.2; DB 3; Length 2505;					
Best Local Similarity 84.7%; Pred. No. 0;					
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;					
QY 48 ATGGGAAATTAATAATCTAGACCAACCAACTTGTACTGATGAATGGAGAAAGATC 107					
Db 1 ATGGGGAATTACAAATCCAGACCAACACAGACTTGTCTGATGAATGGAAGAAAGTT 60					
QY 108 AGTGAATCATATGTTTATCACAATAGAAAGATTAGAGTACCTGCGAGATCAAGGAAAA 167					

Db 1141 GATGAGCAGACATGTTTGATGTGGGCTTACGAGAAAGGACATGATGCCATTTGTTACACTC 1200
Qy 1248 CTGAAGCATTAAGAGACACCAAGATGAATTTGCCCTGTAAATGAATATTTCTCAGCCCTGGA 1307
Db 1201 CTGAAGCACAACAGAGACCCAGAGAGAGTGCCATGTAAACGAATATTTCCAGACCTGGA 1260
Qy 1308 GGAGATGCTCCTATGTGTCTGTTCATACACCTTTGGGGAAGATTTAAAGCATGACAAA 1367
Db 1261 GGAGATGCTCCTATGTGTCTGTTCATACACCTTTGGGGAAGATTTAAAGCATGACAAA 1320
Qy 1368 GAGAGGCGAGATATTTCTCCTCTAAGAGCTGGATTTGCCCTTCAATTTCCATCTCTCAGCTC 1427
Db 1321 GAGAGGCGAGATTTCTCCTCTAAGAGCTGGATTTGCCCTTCAATTTCCATCTCTCAGCTC 1380
Qy 1428 TCAGAAATGAGTTCCATGAGATTAATTTGGCTCAGTTTCTTTGGGAAAGTATATAAAGGA 1487
Db 1381 TCCGAAATCGAGTTCCAGAGATTAATCGGCTCGGGTTCTTTGGGAAAGTCTATAAAGG 1440
Qy 1488 CGATGCGAATAAATAAGTGGCTATAAAGCTTATCGAGCCCAATACCTACTGCTCCAAG 1547
Db 1441 CGATGCGAATAAATAAGTGGCTATAAAGCTTATCGAGCCCAATACCTACTGCTCCAAG 1500
Qy 1548 TCAGATGTGATATGTTTTGCGAGAGGTGTCCATTTCTGCGAGCTCAATCATCCTCTGC 1607
Db 1501 TCAGATGTGATATGTTTTGCGAGAGGTGTCCATTTCTGCGAGCTCAACACCCCTGC 1560
Qy 1608 GTAAATTCAGTTTGGGTGCTTGGTGAATGATCCAGCCAGTTTGCATTTGCACTCAA 1667
Db 1561 GTGTTTCAGTTTGGGTGCTTGGTGAATGATCCAGCCAGTTTGCATTTGCACTCAA 1620
Qy 1668 TACATATCAGGGGTTCTCTGTTCTCCTCTCATGAGCAGAGAGGATTTCTGATTTG 1727
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCTCTCATGAGCAGAGAGATTTCTGATTTG 1680
Qy 1728 CAGTCTAAATTAATTCAGTAGATTTGTCGAAGGATTTGAGTACCTTCAACACCTG 1787
Db 1681 CAGTCTAAATTAATTCAGTAGATTTGTCGAAGGATTTGAGTACCTTCAACACCTG 1740
Qy 1788 ACACAGCAATTAACATGCTGACTTCAACAGTACAAATATTTCTCTATGAGGATGG 1847
Db 1741 ACCAGCCATATACACCGGACCTGAACAGCCCAATATTTCTGCTCTATGAGGATGG 1800
Qy 1848 CATGCTGTGTGGCAGATTTTGGAGATTCAGATTTCTCAGTCTCTGGATGAAGCAAC 1907
Db 1801 CATGCTGTGTGGCAGATTTTGGAGATTCAGATTTCTCAGTCTCTGGATGAAGCAAC 1860
Qy 1908 ATGACAAACAACTGGGACCTCCGTTGGATGGCTTGGGTTTCAAGTGTTCACGATGCACT 1967
Db 1861 ATGACAAACAACTGGGACCTCCGTTGGATGGCTTGGGTTTCAAGTGTTCACGATGCACT 1920
Qy 1968 CGGTACACCATCAAGCAGATGTTCTCAGCTATGCTGTGTCTGTGGGAAATTTCTCACT 2027
Db 1921 AGATACACCATCAAGCAGATGTTCTCAGCTATGCTGTGTGTCTGTGGGAGCTCCTCACT 1980
Qy 2028 GGCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGCGCAGCAGACATGGCTTACAC 2087
Db 1981 GGCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGCGCAGCAGACATGGCTTACAC 2040
Qy 2088 CACATCAGACCTCCATGCTGATTTCCATTTCCAAAGCCCATATCATCTCTGCTGATACGA 2147
Db 2041 CACATCAGACCGGCCATGCGGCTATTCATCCCAAGCCCATATCATCTCTGCTGATACGA 2100
Qy 2148 GGGTGGAAACGATGCTCTGAAGGAAGACCCGAATTTCTGAAGTTCTCATGAAGTTAGAA 2207
Db 2101 GGGTGGAAATGATGCTCTGAAGGAAGACCCGAATTTCTGAAGTTCTCATGAAGTTAGAA 2160
Qy 2208 GAGTGTCTCTGCACATTTGAGCTGATGCTCTCTGATCAATCAAGTAAACAGTGGTCTCTC 2267
Db 2161 GAGTGTCTCTGCACATTTGAGCTGATGCTCTCTGATCAATCAAGTAAACAGTGGTCTCTC 2220
Qy 2268 TCACCTTCTCTCTCTGATTTGCTGGTGAACCGGGGAGGACCTCGCGCGAGTTCATGTTG 2327
Db 2221 TCACCTTCTCTCTCTGATTTGCTGGTGAACCGGGGAGGACCTCGCGCGAGTTCATGTTG 2280

Qy 2328 GCAGCATTAAGAAGTCGTTTTCGAATTTGAATATGCTCTAAATGCAAGTCTTATGCTGCT 2387
Db 2281 GCAGCCTTACGGAGCGGTTTTTGAGTTGGAGTATGCCCTAAATGCAAGTCTTATGCTGGG 2340
Qy 2388 TTGTCCCAAGTGTCTGACAAATATTTCTCTCAAGTCTGTCTTTGGAGGAGATGAAGA 2447
Db 2341 TGTGTCCCAAGTGTGTGAACACACTCTTAATCCGGGCTGTCTTTGGAGGAGATGAATAG 2400
Qy 2448 AGTCTTCAATACACACCCATTGACAAATATGCTATGCTATCCGATCCCATGAGCTCAATG 2507
Db 2401 AGCACCAGATTTCAACTGTTGACAAATACGGCTATGCTGTGATCCCATGAGCTGAGC 2460
Qy 2508 CATTTTCATTTCTGCGAAATATGATAGAGCTTTTGAGGACAGCAGC 2552
Db 2461 CACCTTCACTCCGCGCAAGCAGCAGCAACTTTTGAGGACAGCAAC 2505

RESULT 9
AAD57334 ID AAD57334 standard; cDNA; 2024 BP.
AC AAD57334;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human kinase and phosphatase (KPP-7) cDNA.
XX
KW Human; kinase; phosphatase; KPP; cell proliferative disorder; hepatitis; atherosclerosis; cirrhosis; haemoglobinuria; polycythaemia vera; cancer; psoriasis; thrombocytopaenia; developmental disorder; Reiter's syndrome; renal tubular acidosis; anaemia; mental retardation; Alzheimer's disease; neurological disorder; Parkinson's disease; epilepsy; glomerulonephritis; autoimmune disorder; inflammatory disorder; AIDS; Goodpasture's syndrome; acquired immune deficiency syndrome; immunosuppressive; Crohn's disease; neurotic; transgenic; dermatitis; multiple sclerosis; diabetes mellitus; allergy; gout; Grave's disease; Hashimoto's thyroiditis; bowel syndrome; osteoporosis; rheumatoid arthritis; Sjogren's syndrome; ophthalmological; gene therapy; asthma; anticonvulsant; uropathic; pancreatitis; gene; ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 46..1857
FT /*tag= a
FT /product= "Human KPP"
XX
PN WO2003050084-A2.
XX 19-JUN-2003.
PD
XX 06-DEC-2002; 2002WO-US039126.
PF
XX 07-DEC-2001; 2001US-0340235P.
PR 19-DEC-2001; 2001US-0343007P.
PR 21-DEC-2001; 2001US-0343546P.
PR 04-FEB-2002; 2002US-0354388P.
PR 15-FEB-2002; 2002US-0356757P.
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Kable AE, Chien D, Wilson AD, Swarnakar A, Gorvad AE;
PI Hafalia Aza, Emerling BM, Ramkumar J, Jin P, Griffin JA, Marquis JP;
PI Baughn MR, Chawla NK, Lehr-Mason PM, Khare R, Lee S, Hawkins PR;
PI Becha SD, Lee SY, Sprague WW, Zebbarjadian Y;
XX
DR WPI; 2003-532894/50.
DR P-FSD3; AAE37962.
XX
PT New human kinases and phosphatases and polynucleotides, useful for diagnosing, treating or preventing autoimmune or inflammatory disorders (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis, cancer or hepatitis.


```
QY 1756 TTGCAAGGATGGAGTACCTTACAACTGACAGCAATATACATCGTACTTGA 1815
DB 1754 TTGCAAGGATGGAGTACCTTACAACTGACAGCAATATACATCGTACTTGA 1813
QY 1816 ACAGTCACAATA 1827
DB 1814 ACAGATGCTGTA 1825

RESULT 10
ID ABT09088 standard; DNA; 668 BP.
AC ABT09088;
XX
DT 05-DEC-2002 (first entry)
XX
DE Phase-1 Rat CT gene SEQ ID No 176.
XX
KW Rat; toxicity study; rat toxic response gene; toxicological response;
KW drug development; phase-1 rat CT gene; ds.
XX
OS Rattus sp.
XX
FN WO200266682-A2.
XX
PD 29-AUG-2002.
XX
PF 29-JAN-2002; 2002WO-US002935.
XX
PR 29-JAN-2001; 2001US-0264933P.
PR 26-JUL-2001; 2001US-0308161P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Farris G, Hicken SH, Farr SB;
XX
DR WPI; 2002-674961/72.
XX
PT Evaluating the toxicity of an agent, useful in drug development or in
PT determining toxicological responses to a new drug, by determining the
PT expression of rat toxicologically relevant genes in the test animal in
PT response to the test agent.
XX
PS Disclosure; Page 149; 388pp; English.
XX
CC The invention relates to a method used for evaluating the toxicity of an
CC agent comprising determining the expression of a rat toxic response
CC gene(s) in the test animal in response to the agent. The method is useful
CC in drug development, particularly for conducting toxicity studies and
CC analysis before a new drug or compound is approved for human consumption
CC or use. The method is also useful in determining toxicological responses
CC to a new drug. This polynucleotide sequence represents a phase-1 rat CT
CC gene of the invention
XX
SQ Sequence 668 BP; 162 A; 186 C; 170 G; 149 T; 0 U; 1 Other;

Query Match 15.4%; Score 466.4; DB 6; Length 668;
Best Local Similarity 87.1%; Pred. No. 1.6e-107;
Matches 512; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 1749 GTAGATGTTCCAAAGGATGGAGTACCTTACAACTGACAGCAATATACATCGT 1808
DB 77 GTAGATGTTCCAAAGGATGGAGTACCTTACAACTGACAGCAATATACATCGT 136
QY 1809 GACTTGAACAGTCACAATATTTCTTCTATGAGATGGGCATGCTGTGGCAGATTTT 1868
DB 137 GACCTGAACAGCCACAATATTTCTGCTCTATGAGATGGGCATGCTGTGGCAGATTTT 196
QY 1869 GGAGAAATCAAGATTTTACAGTCTCTGGATGACAGCAATGACAAACAACTGGGAAC 1928
DB 197 GGAGAAATCAAGATTTTCTGAGTCCCTGGATGAAACAAATGACAAAGCAGCCGGGAC 256
```

```
QY 1929 CTCGGTTGGATGGCTCTCTGAGGTGTTTCAGCGAGTGCACCTCGGTACACCATCAAGCAGAT 1988
DB 257 CTGGGCTGGATGGCCCTTGAGGTGTTTCACACAGTGCACGAGATACACCATCAAGGCTGAT 316
QY 1989 GTCTTTCAGTATGCTCTGTGTCTGTGGGAAATTTCTCACTGGCGAAATTCATTTCGCTCAT 2048
DB 317 GTCTTTCAGTTACTCCCTGTGTGTGTGGAGTCTCTCACTGGAGAAATTCATTTCGCTCAT 376
QY 2049 CTCAGCCAGCGGTGGGAGCAGACAGATGGCTTACCACACATCAGACTCCCATTCGC 2108
DB 377 CTCAGCCAGCGGTGGGAGCAGACAGATGGCTTACCACACATCAGACTCCCATTCGC 436
QY 2109 TATTCATATCCCAAGCCCATATCATCTCTGTGATACGAGGGTGGAAACGATGCTCTGAA 2168
DB 437 TATTCATATCCCAAGCCCATCTCATCCCTGTGATACGAGGGTGGAAATGATGCTCTGAA 496
QY 2169 GGAAGACCCGAAATTTCTGAAATGTCATGAAGTGAAGATGAAGATGCTCTGCAACATTGAG 2228
DB 497 GGAGCAGCAGATTTCTGAAATGTCGTTAGCAAACTGGAGAGTGGCTATSCAATGTGGAG 556
QY 2229 CTGATGTCCTCTCATCAAGTAACAGCAGTGGGTCTCTCTCACCTTCTTCTTCTTCTGAT 2288
DB 557 CTATGTCCTCAGCATCAAGTAACAGCAGTGGGTCTCTCTCACCTTCTTCTTCTTCTTCCGAT 616
QY 2289 TGCTGTGTGAACCGGGAGGAGCCTGCGCGAGTCAATGTGGCAGCATTA 2336
DB 617 TGCTGTGTGAACCGGGAGGAGCCTGCGCGAGTCAATGTGGCAGCATTA 664

RESULT 11
AAS33234
ID AAS33234 standard; cDNA; 633 BP.
XX
AC AAS33234;
XX
DT 04-DEC-2001 (first entry)
XX
DE DNA encoding human secreted protein, Seq ID No 193.
XX
KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing; ss.
XX
OS Homo sapiens.
XX
FN WO200155326-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001347.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180828P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
```


CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
CC disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
CC and antibodies can also be used to promote wound healing, maintain organs
CC before transplantation, and support cell culture of primary tissues.

Query Match 15.2%; Score 458.4; DB 4; Length 633;
Best Local Similarity 93.2%; Pred. No. 1.6e-105;
Matches 517; Conservative 5; Mismatches 27; Indels 6; Gaps 4;
QY 956 CTTGAGTGAACAGCTTTTCATAGTGGTTGACCTATGCGAAGAGCATTCAGCTAGTCAA 1015
Db |||||
44 CTTGATGAACAGCNCCTTTTCATAGTGGTTGACCTATGCGAAGAGCATTCAGCTAGTCAA 103
QY 1016 ATTTCTTCTTGATCAGATGTCATTAACATCAACCAAGGAGGATGGGCACACTGG 1075
Db |||||
104 ATTTCTTCTTGATCAGATGTCATTAACATCAACCAAGGAGGATGGGCACACTGG 163
QY 1076 ATTTACATCTGCTTGTTACACAGGTGTCATTCGCTGGTTTCATCTTACTGTAATGG 1135
Db |||||
164 ATTTACATCTGCTTGTTACACAGGTGTCATTCGCTGGTTTCATCTTACTGTAATGG 223
QY 1136 AGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAGATGAGCA 1195
Db |||||
224 AGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAGATGAGCA 283
QY 1196 GACATGTTGATGTTGGCTTATGAAAAGGGCATGATGCCATTTGCACACTCTCTGAAGCA 1255
Db |||||
284 GACATGTTGATGTTGGCTTATGAAAAGGGCATGATGCCATTTGCACACTCTCTGAAGCA 343
QY 1256 TTATAAGAGACCAAGATGATGATGCTGTAATGATATTCAGCTCGGAGAGATGG 1315
Db |||||
344 TTATAAGAGACCAAGATGATGATGCTGTAATGATATTCAGCTCGGAGAGATGG 402
QY 1316 CTCCTATGTTCTCTTCATCACCCTTGGGGAAGATTAAGATCACAAGAGAGAGGCG 1375
Db |||||
403 CTCCTATGTTCTCTTCATCACCCTTGGGGAAGATTAAGATCACAAGAGAGAGGCG 462
QY 1376 AGATATTTCTCTCTTAAGAGCTGATGCTTTCACATTTTC - CATTTTCAGCTCTCAGAA 1433
Db |||||
463 AGATATTTCTCTCTTAAGAGCTGATGCTTTCACATTTTCATCTTCAGCTCTTCAGAAA 522
QY 1434 ATTGAGTTTCCATG - AGATATTTGCTCAGTTCTTTTGGGAAGTATATAA - AGGAGCA 1490
Db |||||
523 TTGGAGTTTCCATGAGGATTTAGGTTCTTTGGGGAAGTATATAAAGGAGCG 582
QY 1491 TGCAGAAATAAATA 1505
Db |||||
583 TGCAGAAATAAATA 597

RESULT 12

AAS33357
ID AAS33357 standard; cDNA; 640 BP.

XX AC AAS33357;

XX DT 04-DEC-2001 (first entry)

XX DE DNA encoding human secreted protein, Seq ID No 316.

XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
XX rheumatoid arthritis; antihypertensive; cardiant; vascular;
XX cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
XX cystostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
XX multiple sclerosis; cancer; hyperproliferative disorder; infection;

KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing; ss.

XX OS Homo sapiens.

XX PN WO200155326-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001347.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225271P.

PR 14-AUG-2000; 2000US-0225275P.

PR 14-AUG-2000; 2000US-0225275P.

PR 14-AUG-2000; 2000US-0225275P.

PR 14-AUG-2000; 2000US-0225275P.

PR 14-AUG-2000; 2000US-0225275P.

PR 14-AUG-2000; 2000US-0225275P.

PR 14-AUG-2000; 2000US-0225275P.

PR 14-AUG-2000; 2000US-0225275P.

PR 14-AUG-2000; 2000US-0225275P.

PR 14-AUG-2000; 2000US-0225275P.

```
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241828P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.

PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451931/48.
DR P-PSDB; AAU20648.
XX New nucleic acids and polypeptides, useful for diagnosing, preventing or
PT treating medical conditions.
XX Claim 1; SEQ ID NO 316; 753pp; English.
XX The invention relates to novel isolated nucleic acid molecules (I)
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits, in
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC the prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression of secreted proteins. (I) and complementary
CC sequences may also be used as DNA probes in diagnostic assays (e.g.
CC polymerase chain reactions (PCR)) to detect and quantitate the presence
CC of similar nucleic acid sequences in samples, and so which patients may
CC be in need of restorative therapy. (II) may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of the secreted
CC proteins. The anti-(II) antibodies and antagonists may also be used to
CC down regulate expression and activity of (II). The anti-(II) antibodies
CC may also be used as diagnostic agents for detecting the presence of (II)
CC in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The
CC disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
CC and antibodies can also be used to promote wound healing, maintain organs
CC before transplantation, and support cell culture of primary tissues.

Query Match 14.3%; Score 432; DB 4; Length 640;
Best Local Similarity 97.9%; Pred. No. 7.8e-99;
Matches 457; Conservative 1; Mismatches 7; Indels 2; Gaps 2;

QY 956 CTTCACTGAACACAGCTTTTCATAGTGTCTTACTAGTGTGCGAAGAGGATGGGACACTGG 1015
DB 44 CTTCACTGAACAGCTTTTCATAGTGTCTTACTAGTGTGCGAAGAGGATGGGACACTGG 103
QY 1016 ATTTCTTCTTGATCAGAATGTCTATAACATCAACCAAGGAGGATGGGACACTGG 1075
DB 104 ATTTCTTCTTGATCAGAATGTCTATAACATCAACCAAGGAGGATGGGACACTGG 163
QY 1076 ATTACACTCTGCTTGCTACCAAGGATGATGCCCTGGTTCAGTCTTCTTACTGATAATGG 1135
DB 164 ATTACACTCTGCTTGCTACCAAGGATGATGCCCTGGTTCAGTCTTCTTACTGATAATGG 223
QY 1136 AGCTGATATGAATCTAGTGGCTTGTGATCCACAGCTCTAGTGGTGAAGAGATGAGCA 1195
DB 224 AGCTGATATGAATCTAGTGGCTTGTGATCCACAGCTCTAGTGGTGAAGAGATGAGCA 283
QY 1196 GACATGTTGATGGGCTTATGAAAAGGGATGATGCCATGTCACACTCTCTGAAGCA 1255
DB 284 GACATGTTGATGGGCTTATGAAAAGGGATGATGCCATGTCACACTCTCTGAAGCA 343
QY 1256 TTATAAGAGACACCAAGATGAATTCGCCCTGTATGAATTTCTCAGCTGGAGGATGG 1315
DB 344 TTATAAGAGACACCAAGATGAATTCGCCCTGTATGAATTTCTCAGCTGGAGGATGG 402
QY 1316 CTCCTATGTCTCTTTCATCACCCTTGGGGAGAGATTAAGAGATGACAAAGAGAGGC 1375
DB 403 CTCCTATGTCTCTTTCATCACCCTTGGGGAGAGATTAAGAGATGACAAAGAGAGGC 462
```

QY 1376 AGATATTCTCTCTCTAA-GAGCTGGATTGCTTCAATTTCCATCTT 1421
 |||||
 Db 463 AGATATTCTCTCTCTAAAGAGCTGGATTGCTTCAATTTCCATCTT 509
 |||||

RESULT 13

AAS80722
 ID AAS80722 standard; cDNA; 1557 BP.

XX AC AAS80722;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #16526.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX ER 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG16535.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits and to assess
 biodiversity.

XX PS Claim 1; SEQ ID NO 16526; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 sequences. (I) is useful as hybridisation probes, polymerase chain
 reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 and in recombinant production of (II). The polynucleotides are also used
 in diagnostics as expressed sequence tags for identifying expressed
 genes. (I) is useful in gene therapy techniques to restore normal
 activity of (II) or to treat disease states involving (II). (II) is
 useful for generating antibodies against it, detecting or quantitating a
 polypeptide in tissue, as molecular weight markers and as a food
 supplement. (II) and its binding partners are useful in medical imaging
 of sites expressing (II). (I) and (II) are useful for treating disorders
 involving aberrant protein expression or biological activity. The
 polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits to assess biodiversity
 and to produce other types of data and products dependent on DNA and
 amino acid sequences. AAS64157-AAS94564 represent novel human diagnostic
 coding sequences of the invention. Note: The sequence data for this
 patent did not appear in the printed specification, but was obtained in
 electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1557 BP; 417 A; 400 C; 459 G; 281 T; 0 U; 0 Other;

Query Match 4.5%; Score 135.6; DB 5; Length 1557;
 Best Local Similarity 68.4%; Pred. No. 1.1e-23;
 Matches 245; Conservative 0; Mismatches 59; Indels 54; Gaps 2;

QY 87 GATGAATGGAAGAAAAAGTCAGTGAATCATATGTTATCACAATAGAAATGCAAGT 146

Db 178 GATGAATGGAAGAAAAAGTCAGTGAATCATATGTTATCACAATAGAAATGCAAGT 237
 |||||
 QY 147 GACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAAGAAATATATTTGGCTCTGATGAA 206
 |||||
 Db 238 GACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAAGAAATATATTTGGTGGTTTGTG 297
 |||||
 QY 207 GCCTT-----CAGTAAA 218
 |||||
 Db 298 ATCTTGTGCTTCCAGGAGTGAAGCTGCAGGCCTTCATGCTGAGTGTACAGCTCATAAA 357
 |||||
 QY 219 GTCAATTTTAATTACCGCACTGAAATGGCTGTCTCTACTTCA-----TTTATGTTGC 272
 |||||
 Db 358 GGCAGTCCGAGCAACCAAGAGTGAGAAAGCAGCAAGATTTATTCAAAAAGCGAAAGAACAA 417
 |||||
 QY 273 ATTTGTGGAGGCAAGAAATCACAATTTTCAACTCTTATGTTGAAAGGGCTCCGCCCATCT 332
 |||||
 Db 418 AGCTCCCAAGGCAAGAAATCACAATTTTCAACTCTTATGTTGAAAGGGCTCCGCCCATCT 477
 |||||
 QY 333 CGACTGACAAAGAAATGGAATTTACAGCCTTGCAATTTAGCAGTTTACAGGATAATGCAG 390
 |||||
 Db 478 CGACTGACAAAGAAATGGAATTTACAGCCTTGCAATTTAGCAGTTTACAGGCGCGAGGAG 535
 |||||

RESULT 14

AAC48526

ID AAC48526 standard; DNA; 2283 BP.

XX AC AAC48526;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 57797.

XX KW Hybridisation assay; genetic mapping; gene expression control;
 protein identification; signal transduction pathway; metabolic pathway;
 promoter; termination sequence; ss.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 23-MAR-1999; 99US-0123548P.

XX PR 25-MAR-1999; 99US-0125788P.

XX PR 29-MAR-1999; 99US-0126264P.

XX PR 01-APR-1999; 99US-0126785P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 21-APR-1999; 99US-0130077P.

XX PR 23-APR-1999; 99US-0130449P.

XX PR 28-APR-1999; 99US-0130891P.

XX PR 30-APR-1999; 99US-0131449P.

XX PR 30-APR-1999; 99US-0132048P.

XX PR 04-MAY-1999; 99US-0132407P.

XX PR 05-MAY-1999; 99US-0132485P.

XX PR 06-MAY-1999; 99US-0132486P.

XX PR 07-MAY-1999; 99US-0132487P.

XX PR 11-MAY-1999; 99US-0132863P.

XX PR 14-MAY-1999; 99US-0134256P.

XX PR 14-MAY-1999; 99US-0134218P.

XX PR 14-MAY-1999; 99US-0134219P.

XX PR 14-MAY-1999; 99US-0134221P.

XX PR 14-MAY-1999; 99US-0134370P.

PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135623P.	PR	03-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137503P.	PR	06-AUG-1999;	99US-0147416P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139453P.	PR	12-AUG-1999;	99US-0148341P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139455P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139456P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139457P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139458P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139461P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139462P.	PR	23-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139463P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139763P.	PR	25-AUG-1999;	99US-0149303P.
PR	21-JUN-1999;	99US-0139817P.	PR	25-AUG-1999;	99US-0149506P.
PR	22-JUN-1999;	99US-0138695P.	PR	26-AUG-1999;	99US-0150884P.
PR	23-JUN-1999;	99US-0140353P.	PR	27-AUG-1999;	99US-0151065P.
PR	23-JUN-1999;	99US-0140354P.	PR	27-AUG-1999;	99US-0151066P.
PR	24-JUN-1999;	99US-0140695P.	PR	30-AUG-1999;	99US-0151080P.
PR	28-JUN-1999;	99US-0140823P.	PR	30-AUG-1999;	99US-0151303P.
PR	29-JUN-1999;	99US-0140991P.	PR	31-AUG-1999;	99US-0151438P.
PR	30-JUN-1999;	99US-0141287P.	PR	01-SEP-1999;	99US-0151930P.
PR	01-JUL-1999;	99US-0141842P.	PR	07-SEP-1999;	99US-0152363P.
PR	02-JUL-1999;	99US-0142055P.	PR	10-SEP-1999;	99US-0153070P.
PR	06-JUL-1999;	99US-0142390P.	PR	13-SEP-1999;	99US-0153758P.
PR	08-JUL-1999;	99US-0142803P.	PR	15-SEP-1999;	99US-0154018P.
PR	09-JUL-1999;	99US-0142920P.	PR	16-SEP-1999;	99US-0154039P.
PR	12-JUL-1999;	99US-0142977P.	PR	20-SEP-1999;	99US-0154779P.
PR	13-JUL-1999;	99US-0143542P.	PR	22-SEP-1999;	99US-0155139P.
PR	14-JUL-1999;	99US-0143624P.	PR	22-SEP-1999;	99US-0155486P.
PR	15-JUL-1999;	99US-0144005P.	PR	24-SEP-1999;	99US-0155659P.
PR	16-JUL-1999;	99US-0144008P.	PR	28-SEP-1999;	99US-0156458P.
PR	18-JUL-1999;	99US-0144086P.	PR	29-SEP-1999;	99US-0156596P.
PR	19-JUL-1999;	99US-0144325P.	PR	04-OCT-1999;	99US-0157117P.
PR	19-JUL-1999;	99US-0144331P.	PR	05-OCT-1999;	99US-0157753P.
PR	19-JUL-1999;	99US-0144332P.	PR	06-OCT-1999;	99US-0157658P.
PR	19-JUL-1999;	99US-0144333P.	PR	07-OCT-1999;	99US-0158029P.
PR	19-JUL-1999;	99US-0144334P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144335P.	PR	12-OCT-1999;	99US-0158369P.
PR	20-JUL-1999;	99US-0144352P.			

us-10-626-173-1.rng

Tue Sep 7 12:10:51 2004

Qy	2103	ATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATACGAGGTGGAAACGCATGT	2162
Db	1504	ATACCAAAAATGCTCATGCAAAACTTTCTGAGCTTCTTCAGAAATGTTGGCAACAGGAG	1563
Qy	2163	CCTGAAGGAAGACCCGAAATTTCTGAAGT	2191
Db	1564	CCTGCTGAAAGACCAGATTTCTCTGAAAT	1592

Search completed: September 6, 2004, 01:05:49
Job time : 771.541 secs

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 5, 2004, 19:33:36 ; Search time 7778.57 Seconds
(without alignments)
16855.624 Million cell updates/sec
Title: US-10-626-173-1
Perfect score: 3025
Sequence: 1 gtcgaccacgctcgccgcccc.....aaaaaaaaagggcgccgc 3025

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues 6940544
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.cm.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pin.*
35: em.htg.rod.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3025	100.0	3025	6	BD269559	Novel CAR
2	3025	100.0	3025	6	AR269111	Sequence
3	3025	100.0	3025	6	AR437545	Sequence
4	2980.4	98.5	3001	9	AF116826	Homo sapi
5	2980	98.5	2980	9	AY303691	Homo sapi
6	2508	82.9	2508	6	AX056437	Sequence
7	2505	82.8	2505	6	BD269560	Novel CAR
8	2505	82.8	2505	6	AR269112	Sequence
9	2505	82.8	2505	6	AR437546	Sequence
10	2010	66.4	3026	6	BD269561	Novel CAR
11	2010	66.4	3026	6	AR269113	Sequence
12	2010	66.4	3026	6	AR437547	Sequence
13	1991.4	65.8	2982	10	AY303692	AY303692 Rattus no
14	1892.2	62.6	2505	6	BD269562	Novel CAR
15	1892.2	62.6	2505	6	AR269114	Sequence
16	1892.2	62.6	2505	6	AR437548	Sequence
17	1730.6	57.2	7607	9	HSN807040	Homo sapi
18	524	17.3	75492	9	EX470253	Human DNA
19	524	17.3	164132	2	AC019284	Homo sapi
20	524	17.3	165947	2	AC015897	Homo sapi
21	466.4	15.4	668	6	AX525654	Sequence
22	199.4	6.6	150272	2	AC073887	Homo sapi
23	199.4	6.6	179582	9	AC093158	Homo sapi
24	196.2	6.5	66066	2	AC026566	Homo sapi
25	177.6	5.9	112358	9	AC105271	Homo sapi
26	166.6	5.5	197553	10	AC124174	Mus muscu
27	165	5.5	269035	2	AC126082	Rattus no
28	165	5.5	276946	2	AC115396	Rattus no
29	153.4	5.1	150272	2	AC073887	Homo sapi
30	153	5.1	146018	2	AC018404	Homo sapi
31	149.4	4.9	66066	2	AC026566	Homo sapi
32	141.2	4.7	178057	10	AC125097	Mus muscu
33	138.2	4.6	258378	2	AC105566	Rattus no
34	110.4	3.6	78331	9	AC098692	Homo sapi
35	107.2	3.5	60037	2	AC100314	Mus muscu
36	107.2	3.5	254819	2	AC144762	Mus muscu
37	106.4	3.5	1672	8	AY096470	Arabidops
38	106.4	3.5	2122	8	AY070086	Arabidops
39	101.4	3.4	28196	9	AC119672	Homo sapi
40	100.2	3.3	1704	6	AX652967	Sequence
41	100.2	3.3	2263	8	AK072690	Oryza sat
42	99.4	3.3	1748	8	AY128938	Arabidops
43	99.4	3.3	2104	8	AY093017	Arabidops
44	98	3.2	1662	6	AX506319	Sequence
45	94.8	3.1	1961	8	AK073747	Oryza sat

ALIGNMENTS

RESULT 1
BD269559
LOCUS
DEFINITION Novel CAR protein and nucleic acid molecules and uses therefor.
ACCESSION BD269559
VERSION BD269559.1 GI:33079327
KEYWORDS JP 2002535962-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3025)
AUTHORS Raju, J.
TITLE Novel CAR protein and nucleic acid molecules and uses therefor
JOURNAL Patent: JP 2002535962-A 1 29-OCT-2002;

COMMENT		MILLENNIUM PHARMACEUTICALS INC	
OS	Homo sapiens (human)		
PN	JP 2002535962-A/1		
PD	29-OCT-2002		
PF	10-DEC-1999 JP 2000586772		
PR	11-DEC-1998 US 60/111938,14-APR-1999 US 09/291839 P1		
JEYASEELAN RAJU			
PC	C12N15/09, C07K16/40, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/10, C12Q1/02,		
PC	C12Q1/68, G01N33/15, G01N33/50, G01N33/68, G01N33/68/A61K38/45,		
PC	A61K39/395,		
PC	A61K45/00, A61P9/04, A61P9/10, A61P35/00, A61P43/00, C12N15/00, PC C12N5/00,		
PC	A61K37/52		
CC	Novel CARP protein and nucleic acid molecules and uses	CC	
therefor			
FH	Key	Location/Qualifiers	
FT	CDS	(48)..(2552).	
FEATURES			
source			
1..3025			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match		100.0%; Score 3025; DB 6; Length 3025;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 3025; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GTGACCCACGCGTCCGCCCTCGAGAAAGAGAACTTATTAATAATGGGAAATTATA	60
DB	1	GTGACCCACGCGTCCGCCCTCGAGAAAGAGAACTTATTAATAATGGGAAATTATA	60
QY	61	AATCTPAGACCAACCAACTTGTACTGATGAATGGAAGAAAGTCAAGTGAATCATATG	120
DB	61	AATCTPAGACCAACCAACTTGTACTGATGAATGGAAGAAAGTCAAGTGAATCATATG	120
QY	121	TTATCACAATAGAAGATTAGAATGACCTCGAGATCAAGAGAAAGAACTGACAGAAC	180
DB	121	TTATCACAATAGAAGATTAGAATGACCTCGAGATCAAGAGAAAGAACTGACAGAAC	180
QY	181	TAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATACCGCCTG	240
DB	181	TAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATACCGCCTG	240
QY	241	AAAATGGGCTGTCTACTTCAATTTATGTTGCATTTGTGGAGGCAAGAAATCACATATTC	300
DB	241	AAAATGGGCTGTCTACTTCAATTTATGTTGCATTTGTGGAGGCAAGAAATCACATATTC	300
QY	301	GAACCTTATGTTGAAGGGCTCGCCCATCTCGACTGACAGAAATGATTTACAGCT	360
DB	301	GAACCTTATGTTGAAGGGCTCGCCCATCTCGACTGACAGAAATGATTTACAGCT	360
QY	361	TGCATTTAGCAGTTTACAAGATAATGCAGAAATGATCACTTCTCTGCTTCACAGTGGAG	420
DB	361	TGCATTTAGCAGTTTACAAGATAATGCAGAAATGATCACTTCTCTGCTTCACAGTGGAG	420
QY	421	CTGATATACAGCAGGTTGGATACCGTGGCTCTACCTGCTTCATATTTGCTACAAATAGCTG	480
DB	421	CTGATATACAGCAGGTTGGATACCGTGGCTCTACCTGCTTCATATTTGCTACAAATAGCTG	480
QY	481	GCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGGATTAATGTCATATTTCAAGATG	540
DB	481	GCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGGATTAATGTCATATTTCAAGATG	540
QY	541	CAGTTTTTTTCACTCCATTCGATATTCAGCGTACTATGGACATGAACAGGTAATTCGCC	600
DB	541	CAGTTTTTTTCACTCCATTCGATATTCAGCGTACTATGGACATGAACAGGTAATTCGCC	600
QY	601	TTCTTTTGAATTTGGTCTGATGTAATGTAAGTGTGAAGTTGAGATAGACCCCTCC	660
DB	601	TTCTTTTGAATTTGGTCTGATGTAATGTAAGTGTGAAGTTGAGATAGACCCCTCC	660
QY	661	ACCTAGCATCTGCARAAGGATTTCTTGAATATTTCRAAACTCTTGTGGAAGAGCAGCA	720
DB	661	ACCTAGCATCTGCARAAGGATTTCTTGAATATTTCRAAACTCTTGTGGAAGAGCAGCA	720
QY	721	AAGCAGATGTGAATGTCTCAAGATAATGAAGACCATGTCCTCCACTTCTCTCGAT	780
DB	721	AAGCAGATGTGAATGTCTCAAGATAATGAAGACCATGTCCTCCACTTCTCTCGAT	780
QY	781	TTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTTGGAACTTCAACCTCATG	840
DB	781	TTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTTGGAACTTCAACCTCATG	840
QY	841	TTGTTAATATCTATGGAGATACCCCTTACACTGGCATGCTTCAATGGCAAAATTTGAAG	900
DB	841	TTGTTAATATCTATGGAGATACCCCTTACACTGGCATGCTTCAATGGCAAAATTTGAAG	900
QY	901	TTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTCA	960
DB	901	TTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTCA	960
QY	961	GTGAAACAGCTTTTTCATAGTGTGTGACCTATGGCAAGAGCATTTGACCTAGTCAAAATTC	1020
DB	961	GTGAAACAGCTTTTTCATAGTGTGTGACCTATGGCAAGAGCATTTGACCTAGTCAAAATTC	1020
QY	1021	TTCTTGATCAGAATGTATATAACATCAACCAAGGAAGGATGGGACACTGGATTAC	1080
DB	1021	TTCTTGATCAGAATGTATATAACATCAACCAAGGAAGGATGGGACACTGGATTAC	1080
QY	1081	ACTCTGCTTCTACCAAGTTCACATTCGCTCGTTCAGTTCTTACTGATTAATGAGCTG	1140
DB	1081	ACTCTGCTTCTACCAAGTTCACATTCGCTCGTTCAGTTCTTACTGATTAATGAGCTG	1140
QY	1141	ATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAAAGATGAGCAGACAT	1200
DB	1141	ATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAAAGATGAGCAGACAT	1200
QY	1201	GTTTGATGGGCTTATGAAAAAGGCATGATGCCATTTGCACACTCTCGAAGCATATA	1260
DB	1201	GTTTGATGGGCTTATGAAAAAGGCATGATGCCATTTGCACACTCTCGAAGCATATA	1260
QY	1261	AGAGACCAAGATGAATTCCTGCTGTAATGAATTTCTCAGCCTCGAGAGATGCTCCT	1320
DB	1261	AGAGACCAAGATGAATTCCTGCTGTAATGAATTTCTCAGCCTCGAGAGATGCTCCT	1320
QY	1321	ATGTGCTGTTCATCACCTTGGGGAAGATTAAAGCATATGACAAAGAGAGGAGATTA	1380
DB	1321	ATGTGCTGTTCATCACCTTGGGGAAGATTAAAGCATATGACAAAGAGAGGAGATTA	1380
QY	1381	TTCTCTCTCTCAAGCTGGATTCCTTCACTTTCAGTCTCTCAGAAATTTGAGT	1440
DB	1381	TTCTCTCTCTCAAGCTGGATTCCTTCACTTTCAGTCTCTCAGAAATTTGAGT	1440
QY	1441	TCCATGAGATTTATGGCTCAGGTTCTTTTGGAAAGTATATAAAGGACATGACGAAATA	1500
DB	1441	TCCATGAGATTTATGGCTCAGGTTCTTTTGGAAAGTATATAAAGGACATGACGAAATA	1500
QY	1501	AAATAGTGGCTATAAACGTTATCGAGCAATACCTACTGCTCCAAAGTCAGATGGGATA	1560
DB	1501	AAATAGTGGCTATAAACGTTATCGAGCAATACCTACTGCTCCAAAGTCAGATGGGATA	1560
QY	1561	TGTTTTGCCGAGAGGTTGCTCATTTCTGCGAGCTCAATCATCTCCTCGCTTAATTCAGTTG	1620
DB	1561	TGTTTTGCCGAGAGGTTGCTCATTTCTGCGAGCTCAATCATCTCCTCGCTTAATTCAGTTG	1620
QY	1621	TGGGTGCTTGTGAATGATCCAGCAGTTTGGCCATTTGCTCACTCAATACATATCAGGGG	1680
DB	1621	TGGGTGCTTGTGAATGATCCAGCAGTTTGGCCATTTGCTCACTCAATACATATCAGGGG	1680
QY	1681	GTTCTCTGTTCTCCCTCTTCATGAGCAAGAGGATTTCTTGATTTGAGCTTAAATTA	1740
DB	1681	GTTCTCTGTTCTCCCTCTTCATGAGCAAGAGGATTTCTTGATTTGAGCTTAAATTA	1740
QY	1741	TTATTCAGTAGATGTTGCCAAAAGGATGAGAGTACCTTCAACCTGACACAGCCCAATTA	1800

1741 TTTATTGAGTATGTTGGCAAGGATGGAGTACCTTCAACAACCTGACACAGCCAAATTA 1800
1801 TACATCGTGAATGACAGTCAACAATTTCTCTATGAGGATGGGATGCTGGTGG 1860
1801 TACATCGTGAATGACAGTCAACAATTTCTCTATGAGGATGGGATGCTGGTGG 1860
1861 CAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAATGACAAACAA 1920
1861 CAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAATGACAAACAA 1920
1921 CTGGGAACCTCCGTTGGATGGCTTCCAGAGTGTTCACGAGTGCATCGGTACACCATCA 1980
1921 CTGGGAACCTCCGTTGGATGGCTTCCAGAGTGTTCACGAGTGCATCGGTACACCATCA 1980
1981 AAGCAGATGTCTTCAAGTATGCTCTGTCTGTGGGAAATTTCTCACTGGCGAAATTTCCAT 2040
1981 AAGCAGATGTCTTCAAGTATGCTCTGTCTGTGGGAAATTTCTCACTGGCGAAATTTCCAT 2040
2041 TCGCTCATCTCAAGCAGCGGCTGGCGGAGCAGACATGGCTTACCAACCATCAGACCTC 2100
2041 TCGCTCATCTCAAGCAGCGGCTGGCGGAGCAGACATGGCTTACCAACCATCAGACCTC 2100
2101 CCATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATACAGGGTGGAAACGAT 2160
2101 CCATTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATACAGGGTGGAAACGAT 2160
2161 GTCTGAAGAAACCCGAAATTTCTGAAGTGTGTATGAAGTGTGAAGAGTGTCTCTGCA 2220
2161 GTCTGAAGAAACCCGAAATTTCTGAAGTGTGTATGAAGTGTGAAGAGTGTCTCTGCA 2220
2221 ACATTGAGCTGATGTCCTCTGATCAAGTAAACAGAGTGGTCTCTCAACCTTCTTCTT 2280
2221 ACATTGAGCTGATGTCCTCTGATCAAGTAAACAGAGTGGTCTCTCAACCTTCTTCTT 2280
2281 CTCTGATTCCTTGGTGAACCCGGGAGGACCTTGGCCGAGTATGATGGCAGCATTAAGAA 2340
2281 CTCTGATTCCTTGGTGAACCCGGGAGGACCTTGGCCGAGTATGATGGCAGCATTAAGAA 2340
2341 GTCGTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTTATGCTGCTTGTCCCAAGTG 2400
2341 GTCGTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTTATGCTGCTTGTCCCAAGTG 2400
2401 CTGACAAATATTCCTCTCAAGTCTGTCTTTGGAGGAGTGAAGAGTCTTCAATACA 2460
2401 CTGACAAATATTCCTCTCAAGTCTGTCTTTGGAGGAGTGAAGAGTCTTCAATACA 2460
2461 CACCCATGACAAATATGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCATTTCT 2520
2461 CACCCATGACAAATATGCTATGATCCGATCCCATGAGCTCAATGCAATTTTCATTTCT 2520
2521 GCCGAAATGATGAGCTTTGAGGACAGCAGTGAAGAGTCTGCGGATATACCTAAGGAG 2580
2521 GCCGAAATGATGAGCTTTGAGGACAGCAGTGAAGAGTCTGCGGATATACCTAAGGAG 2580
2581 AGTTTTTTTCCCGAACTGACAGCAACGATTTCCAAACCGGAGTGGCTTCCAACTATA 2640
2581 AGTTTTTTTCCCGAACTGACAGCAACGATTTCCAAACCGGAGTGGCTTCCAACTATA 2640
2641 ACATTTTACTCTCAAGGCTCTCTTAAATGGGCTTTGTTTACTTGCTCTTAAATTC 2700
2641 ACATTTTACTCTCAAGGCTCTCTTAAATGGGCTTTGTTTACTTGCTCTTAAATTC 2700
2701 CCCACTATTAGCAGGCTTTGGATTTGGCTAAGGAATAATGCAAAAGAACCAAGACA 2760
2701 CCCACTATTAGCAGGCTTTGGATTTGGCTAAGGAATAATGCAAAAGAACCAAGACA 2760
2761 GAATGATATGAAGATTTGTTTAAATTTGTTAAATTTAAATTTAAATTTAGATCGTTACT 2820
2761 GAATGATATGAAGATTTGTTTAAATTTGTTAAATTTAAATTTAAATTTAGATCGTTACT 2820
2821 TGGAAATGGAGCCTAAGTCTCTGCTGACAGATAAATATGTTTCTCGTGGCTGAATT 2880

2821 TGGAAATGGAGCCTAAGTCTCTGCTGACAGATAAATATGTTTCTCGTGGCTGAATT 2880
2881 ATGTAGACTTGTGTTGACAGCTATGGGTTTATTTCTTAGAACATTTGTTCTTTT 2940
2881 ATGTAGACTTGTGTTGACAGCTATGGGTTTATTTCTTAGAACATTTGTTCTTTT 2940
2941 CTCAATATGTTACTTCTAGTGTTCACCTCTGTGATTAAGATTTCTTTGTTGAATAGAAA 3000
2941 CTCAATATGTTACTTCTAGTGTTCACCTCTGTGATTAAGATTTCTTTGTTGAATAGAAA 3000
3001 AAAAAAAGGCGCGC 3025
3001 AAAAAAAGGCGCGC 3025
3001 AAAAAAAGGCGCGC 3025
3001 AAAAAAAGGCGCGC 3025
RESULT 2
AR269111 LOCUS 3025 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6500654.
ACCESSION AR269111
VERSION AR269111.1 GI:29699954
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3025)
AUTHORS Raju,J.
TITLE CARP protein and nucleic acid molecules and uses therefor
JOURNAL Patent: US 6500654-A 1 31-DEC-2002;
FEATURES Location/Qualifiers
source 1..3025
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 3025; DB 6; Length 3025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTGAGCCAGCGCTCCGCGCTTGGAGAGGAGAACTTATAATAAATGGAAATATA 60
Db 1 GTGAGCCAGCGCTCCGCGCTTGGAGAGGAGAACTTATAATAAATGGAAATATA 60
Qy 61 AATCTAGACCAACCCCAAACTTCTGATGAATGGAAGAAAAGTCAATGATATG 120
Db 61 AATCTAGACCAACCCCAAACTTCTGATGAATGGAAGAAAAGTCAATGATATG 120
Qy 121 TTATCAATATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAAGACTGACAGAAC 180
Db 121 TTATCAATATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAAGACTGACAGAAC 180
Qy 181 TAAGGAATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATACCCGACTG 240
Db 181 TAAGGAATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATACCCGACTG 240
Qy 241 AAAATGGCTGTCTCACTTCAATTTATGTTGATTTGGAGGCAAGAAATCACATATTC 300
Db 241 AAAATGGCTGTCTCACTTCAATTTATGTTGATTTGGAGGCAAGAAATCACATATTC 300
Qy 301 GAACTCTTATGTTGAAGGCTCCGCCCATCTCGACTGACAGAAATGGATTACAGCT 360
Db 301 GAACTCTTATGTTGAAGGCTCCGCCCATCTCGACTGACAGAAATGGATTACAGCT 360
Qy 361 TGCATTAGCAGTTTACAAGGATAATGAGAAATGATCACTTCTCTGCTTCAAGTGGAG 420
Db 361 TGCATTAGCAGTTTACAAGGATAATGAGAAATGATCACTTCTCTGCTTCAAGTGGAG 420
Qy 421 CTGATATACAGAGTTGGATAGCGTGGCTCACTGCCCTCCATTTGCTACAAATAGCTG 480
Db 421 CTGATATACAGAGTTGGATAGCGTGGCTCACTGCCCTCCATTTGCTACAAATAGCTG 480
Qy 481 GCCACCTAGAGGCTGCTGATGCTCTTCAACATGGAGCTAATGTCATATTTCAAGATG 540

|||||
1441 TCCATGAGATATATGGCTCAGTCTCTTTTGGGAAGTATATAAGGACGATGAGAAATA 1500
|||||
1501 AAATAGTGGCTATAAAAGCTTATCGAGCCAACTACTGCTCCAACTCAGATGTGGATA 1560
|||||
1501 AAATAGTGGCTATAAAAGCTTATCGAGCCAACTACTGCTCCAACTCAGATGTGGATA 1560
|||||
1561 TCTTTTGGCGAGAGTGCATCTCTGCGAGCTCAATCATCCCTGGTAAATCAGTTTG 1620
|||||
1561 TCTTTTGGCGAGAGTGCATCTCTGCGAGCTCAATCATCCCTGGTAAATCAGTTTG 1620
|||||
1621 TGGGTCCTGCTTGAATGATCCAGCCAGTGTGCAATGTGCACTCAATATCAGGGG 1680
|||||
1621 TGGGTCCTGCTTGAATGATCCAGCCAGTGTGCAATGTGCACTCAATATCAGGGG 1680
|||||
1681 GTTCTCTGTTCTCCCTCTCATGAGCAGAGAGATCTTGATTTGCACTCAATATTA 1740
|||||
1681 GTTCTCTGTTCTCCCTCTCATGAGCAGAGAGATCTTGATTTGCACTCAATATTA 1740
|||||
1741 TTATTTGCAGTAGATGTGCCAAAGCATGGAGTACCTTCAACACTGACACAGCAATTA 1800
|||||
1741 TTATTTGCAGTAGATGTGCCAAAGCATGGAGTACCTTCAACACTGACACAGCAATTA 1800
|||||
1801 TACATCGTGAATGAAAGTCAATATCTTCTATGAGATGGGATGCTGTGGTGG 1860
|||||
1801 TACATCGTGAATGAAAGTCAATATCTTCTATGAGATGGGATGCTGTGGTGG 1860
|||||
1861 CAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAACATCACAAAACAC 1920
|||||
1861 CAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAACATCACAAAACAC 1920
|||||
1921 CTGGAACTCTCGTTGGATGCTCTGAGGTGTTTCAAGCATGCTCGGTACACCATCA 1980
|||||
1921 CTGGAACTCTCGTTGGATGCTCTGAGGTGTTTCAAGCATGCTCGGTACACCATCA 1980
|||||
1981 AAGCAGATGCTTTGAGTATGCTCTGTTGCTGTTGGGAATTTCTACCTGGCGAATTTCCAT 2040
|||||
2041 TCGCTCATCTCAAGCCAGCGCTCGGCAGCAGACATGGCTTACCACCATCAGACCTC 2100
|||||
2041 TCGCTCATCTCAAGCCAGCGCTCGGCAGCAGACATGGCTTACCACCATCAGACCTC 2100
|||||
2101 CCATTGGCTATTCATTCCTCAAGCCATATCATCTCTGCTGATACGAGGTGAAAGCAT 2160
|||||
2101 CCATTGGCTATTCATTCCTCAAGCCATATCATCTCTGCTGATACGAGGTGAAAGCAT 2160
|||||
2161 GTCTTGAAGGAGACCCGAATTTCTGAAGTTGTGATGAAGTTAGAGTGTCTTGCA 2220
|||||
2161 GTCTTGAAGGAGACCCGAATTTCTGAAGTTGTGATGAAGTTAGAGTGTCTTGCA 2220
|||||
2221 ACATTGAGTGTATGCTCTCTGATCAAGTAAACAGAGTGGGTCTCTTCACTTCTTCT 2280
|||||
2221 ACATTGAGTGTATGCTCTCTGATCAAGTAAACAGAGTGGGTCTCTTCACTTCTTCT 2280
|||||
2281 CTTCTGATGCTGCTGTAACCGGAGGACCTGCGGAGTCTGATGCGAGCATTAAGAA 2340
|||||
2281 CTTCTGATGCTGCTGTAACCGGAGGACCTGCGGAGTCTGATGCGAGCATTAAGAA 2340
|||||
2341 GTCGTTTCCGAATTGGAATATGCTCTAAATGCAAGGTCTTATGCTGCTTTGTCCTCAAGTG 2400
|||||
2341 GTCGTTTCCGAATTGGAATATGCTCTAAATGCAAGGTCTTATGCTGCTTTGTCCTCAAGTG 2400
|||||
2401 CTGGCAATATTCCTCAAGGTCTGCTTGGAGGAGATGAAAGAGTCTTCAATACA 2460
|||||
2401 CTGGCAATATTCCTCAAGGTCTGCTTGGAGGAGATGAAAGAGTCTTCAATACA 2460
|||||
2461 CACCCATTGACAAATATGGCTATGATTCGGATCCCATGAGCTCAATGCAATTTTCATCTT 2520
|||||
2461 CACCCATTGACAAATATGGCTATGATTCGGATCCCATGAGCTCAATGCAATTTTCATCTT 2520
|||||
2521 GCCGAAATAGTAGAGCTTTGAGGACAGAGCTGACAGCATTTGGGATATACCTAAGGAG 2580
|||||

Db 2521 GCCGAAATAGTAGAGCTTTGAGGACAGAGCTGACAGCATTCGGCGTATACCTAAGGAG 2580
QY 2581 AGTTTTTTTCCCGAAGTACAGCAACGATTCCAACACCGCAAGCTGGCTTCCAACTATA 2640
Db 2581 AGTTTTTTTCCCGAAGTACAGCAACGATTCCAACACCGCAAGCTGGCTTCCAACTATA 2640
QY 2641 ACATTTTACTCTCAAGAGTCTCCTTAAATGGGCTTGTTTTACTTGTCTCTATTAAATC 2700
Db 2641 ACATTTTACTCTCAAGAGTCTCCTTAAATGGGCTTGTTTTACTTGTCTCTATTAAATC 2700
QY 2701 CCCACTATTAGAGGCTTTGGATTTGCTCTAAGGAATAATATGAAAAAGAACCAAGACA 2760
Db 2701 CCCACTATTAGAGGCTTTGGATTTGCTCTAAGGAATAATATGAAAAAGAACCAAGACA 2760
QY 2761 GAATGTATGAGAAATGTTTTTAAATTTGTAATTTAAATTTGTAATTTAAATTTGTAATTT 2820
Db 2761 GAATGTATGAGAAATGTTTTTAAATTTGTAATTTAAATTTGTAATTTAAATTTGTAATTT 2820
QY 2821 TGGAAATGAGCCCTAAGTCTGTGTGAGCAGATATAATATTTTCTCTGGCTGAAT 2880
Db 2821 TGGAAATGAGCCCTAAGTCTGTGTGAGCAGATATAATATTTTCTCTGGCTGAAT 2880
QY 2881 ATGTAGACTTGTGTTGACAGCTATGGGTTTATTTCTTGAACATTTTCTTTCTTTT 2940
Db 2881 ATGTAGACTTGTGTTGACAGCTATGGGTTTATTTCTTGAACATTTTCTTTCTTTT 2940
QY 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAAGATTTCTTGGTGAATAGAAA 3000
Db 2941 CTCATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAAGATTTCTTGGTGAATAGAAA 3000
QY 3001 AAAAAAAGGCGGCGCGC 3025
Db 3001 AAAAAAAGGCGGCGCGC 3025

RESULT 4
AF116826
LOCUS Homo sapiens clone HH498 putative protein-tyrosine kinase mRNA, PRI 11-MAY-1999
DEFINITION complete cds.
ACCESSION AF116826
VERSION AF116826.1 GI:4768828
KEYWORDS FLI CDNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3001)
AUTHORS Wei,Y.J., Ding,J.F., Xiong,H., Zhou,Y., Hui,R.T. and Liew,C.C.
TITLE Direct Submission
JOURNAL Submitted (25-DEC-1998) Molecular Medicine Center, Cardiovascular Institute and Fu Wai Hospital, 167 Beilishi Road, West District, Beijing 100037, China
FEATURES
Location/Qualifiers
1..3001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HH498"
/tissue_type="heart"
/dev_stage="adult"
52..7259
/codon_start=1
/product="putative protein-tyrosine kinase"
/protein_id="AAD29632.1"
/db_xref="GI:4768828"

CDS

/translation="MGNYKSRFTQCTDEWKKKSVSEVITIERLEDLQIKKEKELTE
LRNIFGSDAFKNVLNRYTENGSLHLCCCTCGGKSHIRTLMLKGRPRLTFNGF
TALHLAVYDKNAELITSLHSGADIQOYVGGTGLTALHATLHLEADVLQHGANY
NIODAVFFPTLHIAAYGHEOVTRLLKFGADVNSGVDRPLHLAAGKFLNIAKL
LMEGSKDVLNAQDNEDHVPJHFCRSRFGHHDIVKYLQSDLEVOHPVNIYGDTPHL
ACYNGKFEVAKBIQISGTESLTKENIFSETAFHSACTYKSIDLVKLLQONVININ
HQGRDGHGLHSACVHGHLRLVQLFDNGADNLVACDPSRSSGKEQDCLMWAYEK

GHDAIVTLKHKRRQDELPCNEYSQPGDGSYVSPFLSKIKSMTKERADILLRA
GLPSHFHLQLSIEFHEIIGSGFKYKGRKRNKIVAIRYANTYCSKSDVMPOR
EVSILCOLNHPVIOFVACGLNDPSQFAIVTQYSGSLFSLHQRILDLQSKLI
AVDVAKMEYLNLTQPIIHRDLNHNILLYEDGHVAVDFESRFLQSLDENDNTQ
PGLNRMAPEVFTQCTRTVIKADVSVALCWEILTGEIIPFALKPAAAAADNAYTHI
RPIGYISIPKPSISLLIRGNWACPSRPEFSEVVMKLEECNLTILMSPASSNSGSL
SPSSSDCLVNRGGPGRSHVALRRFRFELEVALNABSVAAALSQSQAGYSSQGLSLEEM
KSLQTPIDDKTGYVSDPMSHFHSCRNSSSFEDSS"

polyA_signal

2978..2983

ORIGIN

Query Match 98.5%; Score 2980.4; DB 9; Length 3001;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	16	CGGCCCTCGAGAAAGGAACCTTATAATAATGGGAAATATATAAATCTAGACCAACCC	75
Db	20	CTGCCCTGGAGAAAGGAACCTTATAATAATGGGAAATATATAAATCTAGACCAACCC	79
Qy	76	AAACTTGTACTGATGAATGGAAGAAAAAGTCAGTGAATCATATGCTTATCAATAGAAA	135
Db	80	AAACTTGTACTGATGAATGGAAGAAAAAGTCAGTGAATCATATGCTTATCAATAGAAA	139
Qy	136	GATTAGAAGATGACCTGCAGATCAAGGAAAAAGAACTGCAGAACTAAGGAATATATTG	195
Db	140	GATTAGAAGATGACCTGCAGATCAAGGAAAAAGAACTGCAGAACTAAGGAATATATTG	199
Qy	196	GCTCTGATGAAGCCTTCAGTAAAGTCAATTAATTAACCGCACTGAAATGGCTGCTC	255
Db	200	GCTCTGATGAAGCCTTCAGTAAAGTCAATTAATTAACCGCACTGAAATGGCTGCTC	259
Qy	256	TACTTCATTATGCTGCTTGTGAGGAGCAAGAAATACATATTCGAACTCTTATGTTGA	315
Db	260	TACTTCATTATGCTGCTTGTGAGGAGCAAGAAATACATATTCGAACTCTTATGTTGA	319
Qy	316	AAGGCTCCGCCCATCTGCACTGCAAGAAATGGAAATTAACGCTTGCAATTAGCAGTTT	375
Db	320	AAGGCTCCGCCCATCTGCACTGCAAGAAATGGAAATTAACGCTTGCAATTAGCAGTTT	379
Qy	376	ACAGGATATGAGAAATGATCACTTCTGCTTCCAGAGTGGAGCTGATATACAGCAGG	435
Db	380	ACAGGATATGAGAAATGATCACTTCTGCTTCCAGAGTGGAGCTGATATACAGCAGG	439
Qy	436	TTGGATA CGGTGGCCCTCAGTCCCTCCATATGCTACAA TAGCTGGCCACCTAGAGGCTG	495
Db	440	TTGGATA CGGTGGCCCTCAGTCCCTCCATATGCTACAA TAGCTGGCCACCTAGAGGCTG	499
Qy	496	CTGATGCTCTTGGACATGAGCTAAATGTCAA TATTCAGATGCAGTCTTTTTCATC	555
Db	500	CTGATGCTCTTGGACATGAGCTAAATGTCAA TATTCAGATGCAGTCTTTTTCATC	559
Qy	556	CATTGCATATTGACGGTACTATGGACATGAACAGGTAACTCGCCCTCTTTTGAATTTG	615
Db	560	CATTGCATATTGACGGTACTATGGACATGAACAGGTAACTCGCCCTCTTTTGAATTTG	619
Qy	616	GTGCTGATGTAATGTAAGTGGTGAAGTGGAGATAGACCCCTCCACCTAGCATCTGAA	675
Db	620	GTGCTGATGTAATGTAAGTGGTGAAGTGGAGATAGACCCCTCCACCTAGCATCTGAA	679
Qy	676	AAGGATTTTGAATATTGCAAACTCTTGTGGAAGAGGAGCAAGCAGATGTAATG	735
Db	680	AAGGATTTTGAATATTGCAAACTCTTGTGGAAGAGGAGCAAGCAGATGTAATG	739
Qy	736	CTCAAGATAATGAAGCAATGTCCTCCACTCCATTTCTGTTCTCGATTTGGACACCATGATA	795
Db	740	CTCAAGATAATGAAGCAATGTCCTCCACTCCATTTCTGTTCTCGATTTGGACACCATGATA	799
Qy	796	TAGTAAAGTATCTGCTCAAGGATTTGGAAGTTCAACTCATGTTGTTAATCTATG	855
Db	800	TAGTAAAGTATCTGCTCAAGGATTTGGAAGTTCAACTCATGTTGTTAATCTATG	859
Qy	856	GAGATACCCCTTACACTGGCATGTGTAATGGCAAAATTTGAAGTTGCCAAGGAATCA	915

QY	1996	GCTATGCTCTGTGCTGTGGGAAATTTCTCACTGGCGAAATTCATTCGCTCATCTCAAGC	2055	DEFINITION	Homo sapiens cardiac ankyrin repeat kinase mRNA, complete cds.
DB	2000	GCTATGCTCTGTGCTGTGGGAAATTTCTCACTGGCGAAATTCATTCGCTCATCTCAAGC	2059	ACCESSION	AY303691
QY	2056	CAGCGGCTGGCGCAGCAGACATGGCTTACACACACATCAGACCTCCCATTTGGCTATTC	2115	VERSION	AY303691.1 GI:32165605
DB	2060	CAGCGGCTGGCGCAGCAGACATGGCTTACACACACATCAGACCTCCCATTTGGCTATTC	2119	KEYWORDS	
QY	2116	TTCCCAAGCCCATATCATCTCTGCTGATACGAGGCTGGAACGCGATGCTCTGGAAGGAGAC	2175	SOURCE	Homo sapiens (human)
DB	2120	TTCCCAAGCCCATATCATCTCTGCTGATACGAGGCTGGAACGCGATGCTCTGGAAGGAGAC	2179	ORGANISM	Homo sapiens
QY	2176	CGAATTTTCTGAAGTTGTCATGAAGTTGAAGAGTGTCTCTGCAACATTTGAGCTGATGT	2235	REFERENCE	1 (bases 1 to 2980)
DB	2180	CGAATTTTCTGAAGTTGTCATGAAGTTGAAGAGTGTCTCTGCAACATTTGAGCTGATGT	2239	AUTHORS	Jeyaseelan, R.
QY	2236	CTCCTGCATCAAGTAACAGCAGTGGGCTCTCTCACCTTCTCTCTCTCTCTCTCTCTCTCT	2295	TITLE	Cardiac Ankyrin Repeat Kinase (CARK)
DB	2240	CTCCTGCATCAAGTAACAGCAGTGGGCTCTCTCACCTTCTCTCTCTCTCTCTCTCTCTCT	2299	JOURNAL	Unpublished
QY	2296	TGAACCGGGAGGAGCTGCGCGAGTCATGTGGCAGCATTAAGAAGTCGTTTCCGAATTGG	2355	REFERENCE	2 (bases 1 to 2980)
DB	2300	TGAACCGGGAGGAGCTGCGCGAGTCATGTGGCAGCATTAAGAAGTCGTTTCCGAATTGG	2359	AUTHORS	Jeyaseelan, R.
QY	2356	ATATGCTCTAAATGCAAGTCTTATGCTGCTGTTTGTCCCAAGTGTGGCAATATTCCT	2415	TITLE	Direct Submission
DB	2360	ATATGCTCTAAATGCAAGTCTTATGCTGCTGTTTGTCCCAAGTGTGGCAATATTCCT	2419	JOURNAL	Submitted (20-MAY-2003) Cardiovascular Biology, Millennium Pharmaceuticals Inc., 75 Sidney Street, Cambridge, MA 02139, USA
QY	2416	CTCAAGTCTGCTGTTTGGAGGAGATGAAAGAGTCTTCAATACACACCCATTGACAAAT	2475	FEATURES	Location/Qualifiers
DB	2420	CTCAAGTCTGCTGTTTGGAGGAGATGAAAGAGTCTTCAATACACACCCATTGACAAAT	2479	source	1..2980
QY	2476	ATGGCTATGATCCGATCCGATGAGCTCAATGCAATTTTCAATCTTCTGCGGAAATAGTAGCA	2535	CDS	/organism="Homo sapiens"
DB	2480	ATGGCTATGATCCGATCCGATGAGCTCAATGCAATTTTCAATCTTCTGCGGAAATAGTAGCA	2539		/mol_type="mRNA"
QY	2536	GCTTTGAGCAGCAGCAGCTGACAGCATTCGGGCTATCTAAGGAGAGTCTTTTCCCGGAA	2595		/db_xref="taxon:9606"
DB	2540	GCTTTGAGCAGCAGCAGCTGACAGCATTCGGGCTATCTAAGGAGAGTCTTTTCCCGGAA	2599		/chromosome="1"
QY	2596	CTGACAGCAAGATTCGCAACCAACGCGAAGCTGGCTTCCAACTATAACATTTACTCTCAA	2655		/map="1922.3-p13.3"
DB	2600	CTGACAGCAAGATTCGCAACCAACGCGAAGCTGGCTTCCAACTATAACATTTACTCTCAA	2659		31..2538
QY	2656	AGGTCTCTTAAATTTGGCTGTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2715		/note="CARK; protein kinase"
DB	2660	AGGTCTCTTAAATTTGGCTGTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2719		/codon_start=1
QY	2716	CTTTGGATTTGCTTGAAGTAATATGCAAAAGAACCAAGACAGAAATGTATGAAAGA	2775		/product="cardiac ankyrin repeat kinase"
DB	2720	CTTTGGATTTGCTTGAAGTAATATGCAAAAGAACCAAGACAGAAATGTATGAAAGA	2779		/protein_id="AAF72030.1"
QY	2776	ATTGTTTTTAATTTTGAATTAATAAATAAATTTAGATCGTTACTTCTGGAATGGACCTTA	2835		/db_xref="GI:32165606"
DB	2780	ATTGTTTTTAATTTTGAATTAATAAATAAATTTAGATCGTTACTTCTGGAATGGACCTTA	2839		/translation="MGNYKSRPTQCTDEWKKKVSSEYVITIERLEDLQIKELTE
QY	2836	AGTCTGTGGTGGACAGATAATAATTTGTTTCTGCGGCTGAAATTTAGTAGCTTGTGTT	2895		LRI FGSDFAFSKVNLNRTENGLSLLHLCICCGKSHIRTLMLGLRPSRLTRNGF
DB	2840	AGTCTGTGGTGGACAGATAATAATTTGTTTCTGCGGCTGAAATTTAGTAGCTTGTGTT	2899		TALHAYVDNAELITSLHSGADIQGVGGLTALHATLHAGLADVLQHGAV
QY	2896	TGACAGCTATGGGTTATTTCTTAGAACATTTGTTCTTTCTTTCTTTCTTTCTTTCTTTCT	2955		NQDAVFTPLHIAAYYHEQVTRILLKFGADVNVGSDRPLHLASAKFLNIATKL
DB	2900	TGACAGCTATGGGTTATTTCTTAGAACATTTGTTCTTTCTTTCTTTCTTTCTTTCTTTCT	2959		LMEEGSKADNAQDNEDHVPFLHFCRFGHDIVKYLQSDLEQVHVNNITDIPHLH
QY	2956	CTAGTGTTCACTCTGTGATTAAGATTCTTTGGTGAATAG	2997		ACQNGFEVAKETIQISGTESLTKENIFSETAFHSACTYKSIDLVKFLLDONVIN
DB	2960	CTAGTGTTCACTCTGTGATTAAGATTCTTTGGTGAATAG	3001		GHDALVTLLHKYRPODELPCEYSGPGDGSYVSPVPLKIKMTKEKADILLRA
				ORIGIN	EVSLCOLAHNPVIOFVGACLANDPSQALVTQYISGSLFSLHLEOKRILDLQSLKII
				Query Match	98.5%; Score 2980; DB 9; Length 2980;
				Best Local Similarity	100.0%; Pred. No. 0;
				Matches 2980; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
				QY	18 GCCTCGAGAAAGGAGAACTTATAATAAATGGGAAATATAAATCTAGACCAACCCAA 77
				DB	1 GCCTCGAGAAAGGAGAACTTATAATAAATGGGAAATATAAATCTAGACCAACCCAA 60
				QY	78 ACTGTACTGATCAATGGAAGAAAAGTCAGTGAATCATATGTTATCACAATAGAAAGA 137
				DB	61 ACTGTACTGATCAATGGAAGAAAAGTCAGTGAATCATATGTTATCACAATAGAAAGA 120
				QY	138 TTAGAAGATGACCTGCAGATCAAGGAAAAGAACCTGACAGAACTAAGCAATATATTTGGC 197
				DB	121 TTAGAAGATGACCTGCAGATCAAGGAAAAGAACCTGACAGAACTAAGCAATATATTTGGC 180
				QY	198 TCTGATGAAGCCCTTCAAGTAAAGTCAATTTAAATACCACCTGAAAAATGGGCTGTCTCTA 257
				DB	181 TCTGATGAAGCCCTTCAAGTAAAGTCAATTTAAATACCACCTGAAAAATGGGCTGTCTCTA 240
				QY	258 CTTCAATTTATGTTGCAATTTGTGGAGCAAGAAATCACAATTCGAACCTTATGTTGAA 317
				DB	241 CTTCAATTTATGTTGCAATTTGTGGAGCAAGAAATCACAATTCGAACCTTATGTTGAA 300
				QY	318 GGGCTCCGCCCATCTCGACTGACAGAAATGGATTTACAGCCTTCATTTAGCAGTTTAC 377

Db 301 GGGCTCGCCCATCTCGACTGCAAGAAATGGAATTTACAGCCTTGCAATTTAGCAGTTTAC 360
Qy 378 AAGGATAATGAGAAATGATCACTTCTCTGCTTACAGTGGAGCTGATATACAGAGGTT 437
Db 361 AAGGATAATGAGAAATGATCACTTCTCTGCTTACAGTGGAGCTGATATACAGAGGTT 420
Qy 438 GGATAAGGCTGCTCACTGCGCTCCATATTTGCTACAAATAGCTGGCCACCTAGAGCTGCT 497
Db 421 GGATAAGGCTGCTCACTGCGCTCCATATTTGCTACAAATAGCTGGCCACCTAGAGCTGCT 480
Qy 498 GATGCTGTTGCAAGATGAGCTAATGTCATATTTCAAGATGCAATTTTCTCACTCCA 557
Db 481 GATGCTGTTGCAAGATGAGCTAATGTCATATTTCAAGATGCAATTTTCTCACTCCA 540
Qy 558 TTGCATATTCAGCGTACTATGACATGAACAGGTAACCTCGCTTCTTTTGAATTTGGT 617
Db 541 TTGCATATTCAGCGTACTATGACATGAACAGGTAACCTCGCTTCTTTTGAATTTGGT 600
Qy 618 GCTGATGTAATGTAAGTGTGAAGTGGAGTAGACCCCTCCACTAGCATCTGCAAAA 677
Db 601 GCTGATGTAATGTAAGTGTGAAGTGGAGTAGACCCCTCCACTAGCATCTGCAAAA 660
Qy 678 GGATTTGTAATTTGCAAACTCTTGATGGAAGAGGAGCAAAAGCAGATGTGAATGCT 737
Db 661 GGATTTGTAATTTGCAAACTCTTGATGGAAGAGGAGCAAAAGCAGATGTGAATGCT 720
Qy 738 CAAGATAATGAGACCATGCTCCACTCCATTTCTGTTCTGATTTGGAACCATGATATA 797
Db 721 CAAGATAATGAGACCATGCTCCACTCCATTTCTGTTCTGATTTGGAACCATGATATA 780
Qy 798 GTTAAGTATCTGCTGCAAGAGTATTTGGAAGTTCAACCTCATGTTGTTAAATCATATGGA 857
Db 781 GTTAAGTATCTGCTGCAAGAGTATTTGGAAGTTCAACCTCATGTTGTTAAATCATATGGA 840
Qy 858 GATACCCCTTACACCTGGCATGCTCAATGGAATTTGAAAGTTGCCAAGAAATCATC 917
Db 841 GATACCCCTTACACCTGGCATGCTCAATGGAATTTGAAAGTTGCCAAGAAATCATC 900
Qy 918 CAATATACAGAAACAGAAAGTCTGACTAGGAAGAAACATCTTCAGTGAACACGTTTTCAT 977
Db 901 CAATATACAGAAACAGAAAGTCTGACTAGGAAGAAACATCTTCAGTGAACACGTTTTCAT 960
Qy 978 AGTCTGTTGACCTATGCAAGAGCATTTGACCTAGTCAAAATTTCTTCTGATCAGAAATGTC 1037
Db 961 AGTCTGTTGACCTATGCAAGAGCATTTGACCTAGTCAAAATTTCTTCTGATCAGAAATGTC 1020
Qy 1038 ATAAACATCAACCAAGAAAGGATGGGCAACCTGGAATTAACATCTGCTTGTCTACCAAC 1097
Db 1021 ATAAACATCAACCAAGAAAGGATGGGCAACCTGGAATTAACATCTGCTTGTCTACCAAC 1080
Qy 1098 GGTCAATTCGCTGTTGCTAGTTCTTACTGGAATTAAGAGCTGATATGATGCTTGGCT 1157
Db 1081 GGTCAATTCGCTGTTGCTAGTTCTTACTGGAATTAAGAGCTGATATGATGCTTGGCT 1140
Qy 1158 TGTGATCCAGCAGGCTAGTGTGTAAGAAAGATGAGCAGACATGTTTGTATGTTGGCTTAT 1217
Db 1141 TGTGATCCAGCAGGCTAGTGTGTAAGAAAGATGAGCAGACATGTTTGTATGTTGGCTTAT 1200
Qy 1218 GAAAGAGGAGATGATGCAATGTCACCTCTGAAGCATTAAGAGACCAAGATGAA 1277
Db 1201 GAAAGAGGAGATGATGCAATGTCACCTCTGAAGCATTAAGAGACCAAGATGAA 1260
Qy 1278 TTGCTGCTGAATGAATATTTCTACGCTGAGAGAGATGGCTCTATGCTGTTCCATCA 1337
Db 1261 TTGCTGCTGAATGAATATTTCTACGCTGAGAGAGATGGCTCTATGCTGTTCCATCA 1320
Qy 1338 CCTTTGGGAGAGATTAAGAGCATGACAAAAGAGAGAGATTTCTCCTCTAGAGCT 1397
Db 1321 CCTTTGGGAGAGATTAAGAGCATGACAAAAGAGAGAGATTTCTCCTCTAGAGCT 1380
Qy 1398 GGAATGCTTACATTTCCATTTCTAGCTCTCAGAAATGAGTTCCATCAGATTTATGSC 1457
Db 1381 GGAATGCTTACATTTCCATTTCTAGCTCTCAGAAATGAGTTCCATCAGATTTATGSC 1440

Qy 1458 TCAGGTTCTTTTGGGAAAGTATATAAGGACGATGAGAAATAAAATAGTGGCTATAAAA 1517
Db 1441 TCAGGTTCTTTTGGGAAAGTATATAAGGACGATGAGAAATAAAATAGTGGCTATAAAA 1500
Qy 1518 CGTTATCGAGCCAAATACCTACTGCTCCAAAGTCAGATGGGATATGTTTGGCCGAGGCTG 1577
Db 1501 CGTTATCGAGCCAAATACCTACTGCTCCAAAGTCAGATGGGATATGTTTGGCCGAGGCTG 1560
Qy 1578 TCCATTTCTGCGCAGCTCAATCATCCCTCGTAATTTACGTTTGTGGTGTCTTGTCTGAAT 1637
Db 1561 TCCATTTCTGCGCAGCTCAATCATCCCTCGTAATTTACGTTTGTGGTGTCTTGTCTGAAT 1620
Qy 1638 GATCCGAGCAGTTTCCCATTTGCTCACTCAATACATATCAGGGGTTCTCTGTTCTCCCTC 1697
Db 1621 GATCCGAGCAGTTTCCCATTTGCTCACTCAATACATATCAGGGGTTCTCTGTTCTCCCTC 1680
Qy 1698 CTTTCATGAGCAGAGAGGATTTCTTGATTTTGCAGTCTAAATTAATTTATGTCAGTAGATGTT 1757
Db 1681 CTTTCATGAGCAGAGAGGATTTCTTGATTTTGCAGTCTAAATTAATTTATGTCAGTAGATGTT 1740
Qy 1758 GCCAAAGGCATGAGTAGCTTTCACAACTGACACAGCCAAATTTATACATGCTGACTTGAAC 1817
Db 1741 GCCAAAGGCATGAGTAGCTTTCACAACTGACACAGCCAAATTTATACATGCTGACTTGAAC 1800
Qy 1818 AGTCACAAATTTCTTCTCTATGAGGATGGCATGCTGTTGGCAGATTTTGGAGAAATCA 1877
Db 1801 AGTCACAAATTTCTTCTCTATGAGGATGGCATGCTGTTGGCAGATTTTGGAGAAATCA 1860
Qy 1878 AGATTTCTACAGTCTCTGGATGAAGCAACATGACAAAAACAACTGGGAAACCTCGCTTGG 1937
Db 1861 AGATTTCTACAGTCTCTGGATGAAGCAACATGACAAAAACAACTGGGAAACCTCGCTTGG 1920
Qy 1938 ATGGCTCTGAGGTGTTTCAACGATGCACTCGGTACACCATCAAAAGCAGATGCTTTCAGC 1997
Db 1921 ATGGCTCTGAGGTGTTTCAACGATGCACTCGGTACACCATCAAAAGCAGATGCTTTCAGC 1980
Qy 1998 TATGCTCTGCTGCTGCGGAAATTTCTCACTGCGAAATTTCCATTTCCATCTCAAGCCA 2057
Db 1981 TATGCTCTGCTGCTGCGGAAATTTCTCACTGCGGAAATTTCCATTTCCATCTCAAGCCA 2040
Qy 2058 GCGGCTCGGCGACGACATGCTTACCAACCATCAGACCTCCCATTTGGCTTATCCATT 2117
Db 2041 GCGGCTCGGCGACGACATGCTTACCAACCATCAGACCTCCCATTTGGCTTATCCATT 2100
Qy 2118 CCCAAGCCCATATCATCTCTGCTGATACGAGGTTGGAACGATGCTCTGAAGAGACCC 2177
Db 2101 CCCAAGCCCATATCATCTCTGCTGATACGAGGTTGGAACGATGCTCTGAAGAGACCC 2160
Qy 2178 GAATTTTCTGAAAGTGTCTAAGTTAGAAGAGTGTCTCTGCAACATTTGAGCTGATGCT 2237
Db 2161 GAATTTTCTGAAAGTGTCTAAGTTAGAAGAGTGTCTCTGCAACATTTGAGCTGATGCT 2220
Qy 2238 CTTGCTATCAAGTAACAGCAGTGGGTCTCTCACTCTTCTTCTTCTGATTTGGCTGGTG 2297
Db 2221 CTTGCTATCAAGTAACAGCAGTGGGTCTCTCACTCTTCTTCTTCTGATTTGGCTGGTG 2280
Qy 2298 AACCGGAGGAGCCTCGCGGAGTCAATGTCGAGCATTAAAGAGTCTTTTCCAAATTTGAAA 2357
Db 2281 AACCGGAGGAGCCTCGCGGAGTCAATGTCGAGCATTAAAGAGTCTTTTCCAAATTTGAAA 2340
Qy 2358 TATGCTCTAAATCAAGGTCCTATGCTGCTTTTGTCCAAAGTGTGGAACAATTTCTCTCT 2417
Db 2341 TATGCTCTAAATCAAGGTCCTATGCTGCTTTTGTCCAAAGTGTGGAACAATTTCTCTCT 2400
Qy 2418 CAAGGTCGTCTTTGAGGAGATGAAAGAGTCTTCAATACACCCCATTTGCAAAATAT 2477
Db 2401 CAAGGTCGTCTTTGAGGAGATGAAAGAGTCTTCAATACACCCCATTTGCAAAATAT 2460
Qy 2478 GGTATGTATCCCATGCTGCTCAATGCAATTTTCTTCTTGGCAGAAATAGTAGCAGC 2537
Db 2461 GGTATGTATCCCATGCTGCTCAATGCAATTTTCTTCTTGGCAGAAATAGTAGCAGC 2520

QY	2538	TTTGAGGACAGCTGACAGCATTCGGCGTATACCTAAGGAGTGTCTTCCCGAACT	2597	Db	181	AATTTACGCACTGAAATGGCTGTCTCTACTTCAATTTATGTTGCAATTTTGTGGGCAAG	240
Db	2521	TTTGAGGACAGCTGACAGCATTCGGCGTATACCTAAGGAGTGTCTTCCCGAACT	2580	QY	288	AAATCAATATTTGAACTCTTATGTTGAAAGGGCTCGGCCATCTCGACTGACAGAAT	347
QY	2598	GACAGCAACGATTCACCAACCGGCAAGCTGGCTTCCAACTATAACATTTTACTCTCAAAG	2657	Db	241	AAATCAATATTTGAACTCTTATGTTGAAAGGGCTCGGCCATCTCGACTGACAGAAT	300
Db	2581	GACAGCAACGATTCACCAACCGGCAAGCTGGCTTCCAACTATAACATTTTACTCTCAAAG	2640	QY	348	GGATTTACAGCTTTGCAATTTAGCAGTTTACAGGATAATGCAAGATTGATCACTTCTCTG	407
QY	2658	GTCTCTTAATTTGGCTTGTCTTCTTACTTGTCTTATTAATTCACCACTATAGCAGGCT	2717	Db	301	GGATTTACAGCTTTGCAATTTAGCAGTTTACAGGATAATGCAAGATTGATCACTTCTCTG	360
Db	2641	GTCTCTTAATTTGGCTTGTCTTCTTACTTGTCTTATTAATTCACCACTATAGCAGGCT	2700	QY	408	CTTCAAGTGGAGCTGATATACAGCAGTTGGATACGGTGGCTCACTGCCCTCCATATT	467
QY	2718	TTGGATTTGTGCTAAGCAATAATATGCAAAAGAACCAAGACAGAAATGTATGAAGAAT	2777	Db	361	CTTCAAGTGGAGCTGATATACAGCAGTTGGATACGGTGGCTCACTGCCCTCCATATT	420
Db	2701	TTGGATTTGTGCTAAGCAATAATATGCAAAAGAACCAAGACAGAAATGTATGAAGAAT	2760	QY	468	GCTCAATAGCTGCCACCTAGAGGCTGTGATGTCTGTGCAACATCGAGCTAATGTC	527
QY	2778	TGTTTTTAATTTGTAATTTAAAAAATTTAGATCGTTACTTTGAAATGGAGCCTAAG	2837	Db	421	GCTCAATAGCTGCCACCTAGAGGCTGTGATGTCTGTGCAACATCGAGCTAATGTC	480
Db	2761	TGTTTTTAATTTGTAATTTAAAAAATTTAGATCGTTACTTTGAAATGGAGCCTAAG	2820	QY	528	AATATTCAAGATGCAGTTTTTTTTTCACTCCATTCATATTGCAGCGTACTATGGACATGAA	587
QY	2838	TCGTGGTGGACAGATAATATTTTCTCGGCTGAATATGTAGACTTTGTGTTTG	2897	Db	481	AATATTCAAGATGCAGTTTTTTTTTCACTCCATTCATATTGCAGCGTACTATGGACATGAA	540
Db	2821	TCGTGGTGGACAGATAATATTTTCTCGGCTGAATATGTAGACTTTGTGTTTG	2880	QY	588	CAGTAACTCGCCTTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGGA	647
QY	2898	ACAGTATGGGTTTATTTCTTAGAACATTTGTTCAATTTCTTTCTCAATATGTTACTTCT	2957	Db	541	CAGTAACTCGCCTTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGGA	600
Db	2881	ACAGTATGGGTTTATTTCTTAGAACATTTGTTCAATTTCTTTCTCAATATGTTACTTCT	2940	QY	648	GATAGACCTTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAACTCTTTGATG	707
QY	2958	AGTGTTCACCTCTGTGATTAAGATTTCTTTGGTGAATAG	2997	Db	601	GATAGACCTTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAACTCTTTGATG	660
Db	2941	AGTGTTCACCTCTGTGATTAAGATTTCTTTGGTGAATAG	2980	QY	708	GAAGAGCGACAAAGCAGATGTGAATGTCTCAAGATAATGAAGATAATGAAGATAATGAAG	767
RESULT 6							
AX056437							
LOCUS							
AX056437							
DEFINITION							
Sequence 81 from Patent WO0073469.							
ACCESSION							
AX056437.1							
VERSION							
AX056437.1							
KEYWORDS							
Homo sapiens (human)							
SOURCE							
Homo sapiens							
ORGANISM							
Homo sapiens							
REFERENCE							
1							
AUTHORS							
Florman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.							
TITLE							
Protein kinases							
JOURNAL							
Patent: WO 0073469-A 81 07-DEC-2000;							
Sugen, Inc. (US)							
FEATURES							
source							
1..2508							
/organism="Homo sapiens"							
/mol_type="unassigned DNA"							
/db_xref="taxon:9606"							
ORIGIN							
Query Match							
Best Local Similarity 82.9%; Score 2508; DB 6; Length 2508;							
Matches 2508; Conservative 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;							
QY	48	ATGGGAAATTAATAATCTAGACCAACCAACTTGTACTGATGAATGGAAGAAAAAGTC	107	QY	1128	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAAA	1187
Db	1	ATGGGAAATTAATAATCTAGACCAACCAACTTGTACTGATGAATGGAAGAAAAAGTC	60	Db	1081	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAAA	1140
QY	108	AGTGAATCATATGTTATCAATAGAAGATTAGAAGATGACCTGCGAGATCAAGGAAAAA	167	QY	1188	GATGAGCAGACATGTTTGTGCTTATGAAAAGGGCATGATGCCATTTGTCACTC	1247
Db	61	AGTGAATCATATGTTATCAATAGAAGATTAGAAGATGACCTGCGAGATCAAGGAAAAA	120	Db	1141	GATGAGCAGACATGTTTGTGCTTATGAAAAGGGCATGATGCCATTTGTCACTC	1200
QY	168	GAAGTGAAGCACTAAGGAATATTTGGCTCTGATGAAGCTTCAGTAAAGTCAATTTA	227	QY	1248	CTGAAGCATTTAAGAGACCAAGATGAATTTGCCCTGTAATTAATTTCTCAGGCTGGA	1307
Db	121	GAAGTGAAGCACTAAGGAATATTTGGCTCTGATGAAGCTTCAGTAAAGTCAATTTA	180	Db	1201	CTGAAGCATTTAAGAGACCAAGATGAATTTGCCCTGTAATTAATTTCTCAGGCTGGA	1260
QY	228	AATTAACCGACATGAATAATGGGCTGTCTCTACTTCATTTATGTTGCAATTTTGTGGAGCAAG	287	QY	1308	GGAGATGGCTCTGTCTGTCTGTCTCCTTGCATCCCTTGGGGAAGATTTAAGAGCATGACAAA	1367

Db 1361 GGAGATGGCTCCTATGCTGCTGCTTCCATCACCCCTTGGGGAAGATTAAAGCATGACAAA 1320
Qy 1368 GAGAAGCAGATATCTCTCTCTTAAGAGTGGATTGCTTCAATTTCCATCTTCAGCTC 1427
Db 1321 GAGAGGCAGATATCTCTCTCTTAAGAGTGGATTGCTTCAATTTCCATCTTCAGCTC 1380
Qy 1428 TCAGAAATTCAGTTCATGAGATATATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487
Db 1381 TCAGAAATTCAGTTCATGAGATATATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
Qy 1488 CGATGCAAAATAAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCCAAG 1547
Db 1441 CGATGCAAAATAAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCCAAG 1500
Qy 1548 TCAGATGAGATATGTTTCCGAGAGGTGTCATTTCTCTGCGAGCTCAATCATCCCTGC 1607
Db 1501 TCAGATGAGATATGTTTCCGAGAGGTGTCATTTCTCTGCGAGCTCAATCATCCCTGC 1560
Qy 1608 GTAATTCAGTTTGGGTGCTTCTCTGATGATCCAGCCAGTTTGCCATTGCTACTCAA 1667
Db 1561 GTAATTCAGTTTGGGTGCTTCTCTGATGATCCAGCCAGTTTGCCATTGCTACTCAA 1620
Qy 1668 TACATATCAGGGGTCTCTGTTCTCTCCCTCTCATGAGCAGAAGGATCTTGATTG 1727
Db 1621 TACATATCAGGGGTCTCTGTTCTCTCCCTCTCATGAGCAGAAGGATCTTGATTG 1680
Qy 1728 CAGTCTAAATTAATTTAGTTCAGTAGATGTTGCCAAGGCATGGATGCTTACAACTG 1787
Db 1681 CAGTCTAAATTAATTTAGTTCAGTAGATGTTGCCAAGGCATGGATGCTTACAACTG 1740
Qy 1788 ACAGAGCAATATATACATCTGACTTGAACAGTCACAATATTTCTCTATGAGGATGG 1847
Db 1741 ACAGAGCAATATATACATCTGACTTGAACAGTCACAATATTTCTCTATGAGGATGG 1800
Qy 1848 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGATGAAGACAA 1907
Db 1801 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGATGAAGACAA 1860
Qy 1908 ATGCAAAAACACTGGGACCTCCGTTGGATGGCTCCTGAGTGTTCACGAGTGCAT 1967
Db 1861 ATGCAAAAACACTGGGACCTCCGTTGGATGGCTCCTGAGTGTTCACGAGTGCAT 1920
Qy 1968 CGGTACACCATCAAGCAGATGCTTCAGCTATGCTCTGTGTCTGTGGAAATTTCTCAT 2027
Db 1921 CGGTACACCATCAAGCAGATGCTTCAGCTATGCTCTGTGTCTGTGGAAATTTCTCAT 1980
Qy 2028 GGGCAAAATTCATTCGCTCATCTCAAGCCAGCGCTCGGCGACGACATGCTTACCAC 2087
Db 1981 GGGCAAAATTCATTCGCTCATCTCAAGCCAGCGCTCGGCGACGACATGCTTACCAC 2040
Qy 2088 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATACA 2147
Db 2041 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGCTGATACA 2100
Qy 2148 GGGTGGAAACGATGCTCTGAAGAGAGCCGAAATTTCTGAAGTTGTCTCATGAAGTAGNA 2207
Db 2101 GGGTGGAAACGATGCTCTGAAGAGAGCCGAAATTTCTGAAGTTGTCTCATGAAGTAGNA 2160
Qy 2208 GAGTGTCTCTGCAACATTCAGTGTCTCTCTGCAATCAAGTAAACAGCAGTGGGTCTCTC 2267
Db 2161 GAGTGTCTCTGCAACATTCAGTGTGTCTCTCTGCAATCAAGTAAACAGCAGTGGGTCTCTC 2220
Qy 2268 TCACTTCTTCTTCTCTGATGCTGCTGTAACCGGGAGGACCTCGGCGGAGTCAATG 2327
Db 2221 TCACTTCTTCTTCTCTGATGCTGCTGTAACCGGGAGGACCTCGGCGGAGTCAATG 2280
Qy 2328 GCACATTAAGAGTCTGTTTCGAATTCGAATATGCTCTAAATTCGAAGTCCCTATGCTGCT 2387
Db 2281 GCACATTAAGAGTCTGTTTCGAATTCGAATATGCTCTAAATTCGAAGTCCCTATGCTGCT 2340
Qy 2388 TTGTCCCAAGTCTGGACAAATATTCCTCTCAAGGTCGTGTTTGGAGGAGTGAAGA 2447
Db 2341 TTGTCCCAAGTCTGGACAAATATTCCTCTCAAGGTCGTGTTTGGAGGAGTGAAGA 2400

Qy 2448 AGTCTTCAATACACACCCATTGACAAAATATGCTATGATCCGATCCCATGAGCTCAATG 2507
Db 2401 AGTCTTCAATACACACCCATTGACAAAATATGCTATGATCCGATCCCATGAGCTCAATG 2460
Qy 2508 CATTTTCATCTTTCGGAATATAGTAGCAGCTTTGAGGACAGCAGCTGA 2555
Db 2461 CATTTTCATCTTTCGGAATATAGTAGCAGCTTTGAGGACAGCAGCTGA 2508

RESULT 7
BD269560 2505 bp DNA linear PAT 17-JUL-2003
LOCUS Novel CARL protein and nucleic acid molecules and uses therefor.
DEFINITION
ACCESSION BD269560.1 GI:33079328
VERSION JP 2002535962-A/2.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2505)
Raju,J.
AUTHORS
TITLE Novel CARL protein and nucleic acid molecules and uses therefor
JOURNAL MILLENNIUM PHARMACEUTICALS INC
COMMENT
OS Homo sapiens (human)
PN JP 2002535962-A/2
PD 29-OCT-2002
PF 10-DEC-1999 JP 2000586772
PR 11-DEC-1998 US 60/111938, 14-APR-1999 US 09/291839 PI
JEVASEELAN RAJU
PC C12N15/09, C07K16/40, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/10, C12Q1/02.
PC C12Q1/58, G01N33/15, G01N33/50, G01N33/53, G01N33/68//A61K38/45,
PC A61K39/395,
PC A61K45/00, A61P9/04, A61P9/10, A61P95/00, A61P43/00, C12N15/00, PC
C12N5/00,
PC A61K37/52
CC Novel CARL protein and nucleic acid molecules and uses CC
therefor
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..2505
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 82.8%; Score 2505; DB 6; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 48 ATGGGAATTTATAAATCTAGACCAACCCAACTGTGCTGATGATGGAAGAAAAAGTC 107
Db 1 ATGGGAATTTATAAATCTAGACCAACCCAACTGTGCTGATGATGGAAGAAAAAGTC 60
Qy 108 AGTGAATCATATGTTATCAATATAGAAAGATTAGAAAGATGACCTGCAGATCAAGGAAAA 167
Db 61 AGTGAATCATATGTTATCAATATAGAAAGATTAGAAAGATGACCTGCAGATCAAGGAAAA 120
Qy 168 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAGCCCTTCAGTAAAGTCAATTTA 227
Db 121 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAGCCCTTCAGTAAAGTCAATTTA 180
Qy 228 AATTACCGCACTGAAATGGGCTGTCTTACTTCAATTTATGTTGATTTGTGGAGCAAG 287
Db 181 AATTACCGCACTGAAATGGGCTGTCTTACTTCAATTTATGTTGATTTGTGGAGCAAG 240
Qy 288 AATATCATATTCGAATCTTATGTTGAAAGGGCTCCGCCCATCTCGACTGACAAGAAAT 347
Db 241 AATATCATATTCGAATCTTATGTTGAAAGGGCTCCGCCCATCTCGACTGACAAGAAAT 300

Qy	348	GGATTTACAGCCTTGCAATTTAGCAGTTTACAAGATAAATCGAGAAATTGATCACTTCTCTG	407
Db	301	GGATTTACAGCCTTGCAATTTAGCAGTTTACAAGAGTAATCGAGAAATTGATCACTTCTCTG	360
Qy	408	CTTCACAGTGGAGCTGATATACACAGGTTGGATACGGTGGCTCACTGCCCTCCATATT	467
Db	361	CTTCACAGTGGAGCTGATATACACAGGTTGGATACGGTGGCTCACTGCCCTCCATATT	420
Qy	468	GCTACAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTGCAACATGGAGCTAATGTC	527
Db	421	GCTACAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTGCAACATGGAGCTAATGTC	480
Qy	528	AATATTCAAGATGCAAGTTTTTTTTCACTCCATTCGATATTGGACGCTACTATGGACATGAA	587
Db	481	AATATTCAAGATGCAAGTTTTTTTTCACTCCATTCGATATTGGACGCTACTATGGACATGAA	540
Qy	588	CAGTAACTCGCCTTCTTTTGAATTTTGGTGCTGATGTAATGTAAGTGGTGAAGTTGGGA	647
Db	541	CAGTAACTCGCCTTCTTTTGAATTTTGGTGCTGATGTAATGTAAGTGGTGAAGTTGGGA	600
Qy	648	GATAGACCCCTCCACTAGCATCTGC AAAAGGATTTCTTGAATATTGGCAAAACTCTTGATG	707
Db	601	GATAGACCCCTCCACTAGCATCTGC AAAAGGATTTCTTGAATATTGGCAAAACTCTTGATG	660
Qy	708	GAAAGGCGAGCAAAAGCAGATGTGAATGCTCAAGATAAATGAAGA CATGTGCCACTCCAT	767
Db	661	GAAAGGCGAGCAAAAGCAGATGTGAATGCTCAAGATAAATGAAGA CATGTGCCACTCCAT	720
Qy	768	TTCTGTTCTCGATTTGGACACCATGATATAGTTTAAGTATCTGCTGCAAGTGAATTGGAA	827
Db	721	TTCTGTTCTCGATTTGGACACCATGATATAGTTTAAGTATCTGCTGCAAGTGAATTGGAA	780
Qy	828	GTTCAACCTCATGTGTTTAATATCTATGGAGATACCCCTTACACTGGCATGCTACAAT	887
Db	781	GTTCAACCTCATGTGTTTAATATCTATGGAGATACCCCTTACACTGGCATGCTACAAT	840
Qy	888	GGCAAAATTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACGAAAGTCTGACTAAG	947
Db	841	GGCAAAATTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACGAAAGTCTGACTAAG	900
Qy	948	GAACAACCTTCAGTGAACAGCTTTTTCATAGTGCCTGTACTATGGCAAGAGATTTGAC	1007
Db	901	GAACAACCTTCAGTGAACAGCTTTTTCATAGTGCCTGTACTATGGCAAGAGATTTGAC	960
Qy	1008	CTAGTCAAAATTTCTTTGATCAGAATGTCATAAA CATCAACCAAGGAAGGATGGG	1067
Db	961	CTAGTCAAAATTTCTTTGATCAGAATGTCATAAA CATCAACCAAGGAAGGATGGG	1020
Qy	1068	CACACTGGATTACACTCTGCTTGTACACGGTCAATTCGGCTGGTTCAGTTCTTACTG	1127
Db	1021	CACACTGGATTACACTCTGCTTGTACACGGTCAATTCGGCTGGTTCAGTTCTTACTG	1080
Qy	1128	GATAATGGAGCTGATGAAATCTAGTGGCTTTGATCCCAAGAGCTTAGTGGTGA AAAA	1187
Db	1081	GATAATGGAGCTGATGAAATCTAGTGGCTTTGATCCCAAGAGCTTAGTGGTGA AAAA	1140
Qy	1188	GATGAGCAGACATGTTTTCATGTGGGCTTTA GAAAAAGGCGATGATGCCATTGTCCACTC	1247
Db	1141	GATGAGCAGACATGTTTTCATGTGGGCTTTA GAAAAAGGCGATGATGCCATTGTCCACTC	1200
Qy	1248	CTGAAGCAATTATAAGAGACCAAGATGAATTCGCCCTGTAATGAATATTTTCAGCCTGGA	1307
Db	1201	CTGAAGCAATTATAAGAGACCAAGATGAATTCGCCCTGTAATGAATATTTTCAGCCTGGA	1260
Qy	1308	GGAGATGCTCCTATGTCTGTGTTCCATCACCCCTTGGGGAAGATTTAAAGCATGACAAAA	1367
Db	1261	GGAGATGCTCCTATGTCTGTGTTCCATCACCCCTTGGGGAAGATTTAAAGCATGACAAAA	1320
Qy	1368	GAGAAAGGAGATATTCTCCTTAAGAGCTGGATTGCCCTTACATTTCCATTTTCAGCTC	1427
Db	1321	GAGAAAGGAGATATTCTCCTTAAGAGCTGGATTGCCCTTACATTTCCATTTTCAGCTC	1380

Qy	1428	TCGAAATTCAGTTCCTCCTGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAAGA	1488
Db	1381	TCGAAATTCAGTTCCTCCTGAGATTTATGGCTCAGGTTCTTTTGGGAAAGTATATAAAGA	1440
Qy	1488	CGATGCAGAAATAAAATAGTGGCTATAAAACGTTATCGAGGCCAATACCTACTGCTCCAAG	1547
Db	1441	CGATGCAGAAATAAAATAGTGGCTATAAAACGTTATCGAGGCCAATACCTACTGCTCCAAG	1500
Qy	1548	TCAGATCTGATATGTTTTGCGAGAGGTGTCCATTTCTCTGCGCAGCTCAATATCCTCTGC	1607
Db	1501	TCAGATCTGATATGTTTTGCGAGAGGTGTCCATTTCTCTGCGCAGCTCAATATCCTCTGC	1560
Qy	1608	GTAATTCAGTTTCTGGGTGCTTGCTTGAATGATCCAGGCCAGTTTTCGCCATTCCTCAAA	1667
Db	1561	GTAATTCAGTTTCTGGGTGCTTGCTTGAATGATCCAGGCCAGTTTTCGCCATTCCTCAAA	1620
Qy	1668	TACATATCAGGGGGTCTCTTGTTCTCCCTCTTTCATGAGCAGAAAGGATTTCTTGATTTG	1727
Db	1621	TACATATCAGGGGGTCTCTTGTTCTCCCTCTTTCATGAGCAGAAAGGATTTCTTGATTTG	1680
Qy	1728	CAGTCTAAATTAATTTATGAGTAGATGTTTGCAGAGGATGAGATACCTTCACACCTG	1787
Db	1681	CAGTCTAAATTAATTTATGAGTAGATGTTTGCAGAGGATGAGATACCTTCACACCTG	1740
Qy	1788	ACAGGCCAATTAACATCGTGACTTTGAACAGTCAACAATATTTCTCTATCAGGATGGG	1847
Db	1741	ACAGGCCAATTAACATCGTGACTTTGAACAGTCAACAATATTTCTCTATCAGGATGGG	1800
Qy	1848	CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTCTAGTCTCTGGATGAAGACAAC	1907
Db	1801	CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTCTAGTCTCTGGATGAAGACAAC	1860
Qy	1908	ATGACAAAACAACTGGGACCTTCGATGGCTCCTGAGTGTTTACGAGTGCACT	1967
Db	1861	ATGACAAAACAACTGGGACCTTCGATGGCTCCTGAGTGTTTACGAGTGCACT	1920
Qy	1968	CGGTACACCATCAAGCAGATGTCCTCAGCTATGCTCTGTGCTGTGGAAATTCCTACT	2027
Db	1921	CGGTACACCATCAAGCAGATGTCCTCAGCTATGCTCTGTGCTGTGGAAATTCCTACT	1980
Qy	2028	GGCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGCGACGACAGATGCTTACCAC	2087
Db	1981	GGCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGCGACGACAGATGCTTACCAC	2040
Qy	2088	CACATCAGACCTCCCATTTGGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATACGA	2147
Db	2041	CACATCAGACCTCCCATTTGGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATACGA	2100
Qy	2148	GGGTGGAACGCATGCTCTGAAGGAAGACCCGAATTTTCTGAAGTTGTCTATGAAGTTAGAA	2207
Db	2101	GGGTGGAACGCATGCTCTGAAGGAAGACCCGAATTTTCTGAAGTTGTCTATGAAGTTAGAA	2160
Qy	2208	GAGTGTCTCTGCAACATTTGAGCTGATGTCCTCTGCATCAAGTAAACAGCAGTGGGCTCTC	2267
Db	2161	GAGTGTCTCTGCAACATTTGAGCTGATGTCCTCTGCATCAAGTAAACAGCAGTGGGCTCTC	2220
Qy	2268	TCACCTCTTCTTCTCTGATTCCTGTGGTGAACCCGGGAGGACCTGGCCGGAGTCAATG	2327
Db	2221	TCACCTCTTCTTCTCTGATTCCTGTGGTGAACCCGGGAGGACCTGGCCGGAGTCAATG	2280
Qy	2328	GCAGCATTAAGAGTCTGTTTCGAAATTTGGAATATGCTCTAAATGCAAGGCTCTATGCTGCT	2387
Db	2281	GCAGCATTAAGAGTCTGTTTCGAAATTTGGAATATGCTCTAAATGCAAGGCTCTATGCTGCT	2340
Qy	2388	TTGTCCTCAAGTCTCGCAATAATTTCTCTCAAGGCTCTCTCTTTTGGAGGAGATGAAAAGA	2447
Db	2341	TTGTCCTCAAGTCTCGCAATAATTTCTCTCAAGGCTCTCTCTTTTGGAGGAGATGAAAAGA	2400
Qy	2448	AGTCTTCAATACACACCCATGACAAATATGGCTATGATCCGATCCCATGAGTCAATG	2507
Db	2401	AGTCTTCAATACACACCCATGACAAATATGGCTATGATCCGATCCCATGAGTCAATG	2460
Qy	2508	CATTTTCAATTTCTTGGCGAAATAGTAGCAGCTTTGAGGACAGCAGC	2552

Db 2461 CATTTTCATCTTCCGCAATAGTAGCAGCTTTGAGGACGACG 2505

RESULT 8
LOCUS AR269112 2505 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 3 from patent US 6500654.
ACCESSION AR269112
VERSION AR269112.1 GI:29699955
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2505)
AUTHORS Raju J.
TITLE CARP protein and nucleic acid molecules and uses therefor
JOURNAL Patent: US 6500654-A 3 31-DEC-2002;
FEATURES
source
1. 2505
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 82.8%; Score 2505; DB 6; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 ATGGGAATTAATAATCTAGACCAACCCAACTGTACTGATGAATGGAGAAAAAGTC 107
Db 1 ATGGGAATTAATAATCTAGACCAACCCAACTGTACTGATGAATGGAGAAAAAGTC 60

Qy 108 AGTGAATCATATGTTATFACAATAGAAAGATTAGAAGATGACCTGCAGATCAAGAAAAA 167
Db 61 AGTGAATCATATGTTATFACAATAGAAAGATTAGAAGATGACCTGCAGATCAAGAAAAA 120

Qy 168 GAATGACAGACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 227
Db 121 GAATGACAGACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180

Qy 228 AATTACCGCACTGAAATGGCTGTCTCTACTTCAITTTATGTTGCAATTTGTGGAGCAAG 287
Db 181 AATTACCGCACTGAAATGGCTGTCTCTACTTCAITTTATGTTGCAATTTGTGGAGCAAG 240

Qy 288 AATFACATATTCGAACTTTATGTTGAAAGGCTCGGCCATCTCGACTGACAGAAAT 347
Db 241 AATFACATATTCGAACTTTATGTTGAAAGGCTCGGCCATCTCGACTGACAGAAAT 300

Qy 348 GGATTTACAGCTTTCATTTAGCAGTTTACAAGGATAATGCAGAAATTGATCACTTCTCTG 407
Db 301 GGATTTACAGCTTTCATTTAGCAGTTTACAAGGATAATGCAGAAATTGATCACTTCTCTG 360

Qy 408 CTTTCACAGTGGAGTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 467
Db 361 CTTTCACAGTGGAGTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 420

Qy 468 GCTACATAGCTGCCACCTAGAGGCTGCTGATGCTGCTGCAATGGAGCTAATGTC 527
Db 421 GCTACATAGCTGCCACCTAGAGGCTGCTGATGCTGCTGCAATGGAGCTAATGTC 480

Qy 528 AATATTCAGAGTGCAGTCTTTTTCCTCCATATTCATATTCGACGCTACTATGGACATGAA 587
Db 481 AATATTCAGAGTGCAGTCTTTTTCCTCCATATTCATATTCGACGCTACTATGGACATGAA 540

Qy 588 CAGTAACTGGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGA 647
Db 541 CAGTAACTGGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGA 600

Qy 648 GATAGACCTTCCACCTAGCATCTGCAAAAAGGATCTTGAATATTCGAAAACTCTTGATG 707
Db 601 GATAGACCTTCCACCTAGCATCTGCAAAAAGGATCTTGAATATTCGAAAACTCTTGATG 660

Qy 708 GAAGAGGCGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTGCCACTCCAT 767

Db 661 GAAGAGGCGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720

Qy 768 TTCTGTTCTCGATTTGGACACCAATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTG 827

Db 721 TTCTGTTCTCGATTTGGACACCAATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTG 780

Qy 828 GTTCAACCTCATGTTGTAAATATCTATGAGATACCCCTTACACCTGCGCATGCTACAAAT 887

Db 781 GTTCAACCTCATGTTGTAAATATCTATGAGATACCCCTTACACCTGCGCATGCTACAAAT 840

Qy 888 GGCAAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947

Db 841 GGCAAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900

Qy 948 GAAACATCTTCAGTGAACACAGCTTTTTCATAGTGTGTACCTATGGCAAGAGCATTTGAC 1007

Db 901 GAAACATCTTCAGTGAACACAGCTTTTTCATAGTGTGTGTACCTATGGCAAGAGCATTTGAC 960

Qy 1008 CTAGTCAAAATTTCTTCTTGATCAGAAATGTCAATAACATCAACCAAGGAAGGATGGG 1067

Db 961 CTAGTCAAAATTTCTTCTTGATCAGAAATGTCAATAACATCAACCAAGGAAGGATGGG 1020

Qy 1068 CACTGGAATACACTCTGCTTGTACCAAGGTCAATTCGGCTGGTTCAGTCTTACTG 1127

Db 1021 CACTGGAATACACTCTGCTTGTACCAAGGTCAATTCGGCTGGTTCAGTCTTACTG 1080

Qy 1128 GATAATGGAGCTCATATGAATCTAGTGGCTTGTGATCCCGAGAGGTCTAGTGGTGAATA 1187

Db 1081 GATAATGGAGCTCATATGAATCTAGTGGCTTGTGATCCCGAGAGGTCTAGTGGTGAATA 1140

Qy 1188 GATGAGCAGACATGTTTGTAGTGGCTTATGAAAAAGGCGATGATGCCATTTGTCACCTC 1247

Db 1141 GATGAGCAGACATGTTTGTAGTGGCTTATGAAAAAGGCGATGATGCCATTTGTCACCTC 1200

Qy 1248 CTGAAGCATTTAAGAGACCAAGATGAATGGCCCTGTAATGAATATTTCTCAGGCTGA 1307

Db 1201 CTGAAGCATTTAAGAGACCAAGATGAATGGCCCTGTAATGAATATTTCTCAGGCTGA 1260

Qy 1308 GAGATGGCTCTCTATGCTGTTCATCACCCCTTGGGGAAGATTAAAAAGCATGACAAAA 1367

Db 1261 GAGATGGCTCTCTATGCTGTTCATCACCCCTTGGGGAAGATTAAAAAGCATGACAAAA 1320

Qy 1368 GAGAGGCGATATTTCTCTCTTAAGAGCTGATTTGCCCTTCACATTTCCATCTTCAGTCT 1427

Db 1321 GAGAGGCGATATTTCTCTCTTAAGAGCTGATTTGCCCTTCACATTTCCATCTTCAGTCT 1380

Qy 1428 TCAGAAATTCAGTTCATGAGATTATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487

Db 1381 TCAGAAATTCAGTTCATGAGATTATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440

Qy 1488 CGATGCAAAATTAATATAGTGGCTATAAAAAGTTATCGAGCAATACCTTACTGCTCCAG 1547

Db 1441 CGATGCAAAATTAATATAGTGGCTATAAAAAGTTATCGAGCAATACCTTACTGCTCCAG 1500

Qy 1548 TCAGATGCGATATGTTTTCGAGAGGTTGTCATTTCTGCGAGCTCAATCATCCCTGTC 1607

Db 1501 TCAGATGCGATATGTTTTCGCGAGAGGTTGTCATTTCTGCGAGCTCAATCATCCCTGTC 1560

Qy 1608 GTAATTCAGTTTCTGGTGTCTTGTGAATGATPCCAGCCAGTTTGGCAATTTCTCACTCAA 1667

Db 1561 GTAATTCAGTTTCTGGTGTCTTGTGAATGATPCCAGCCAGTTTGGCAATTTCTCACTCAA 1620

Qy 1668 TACATATCAGGGGTTCTGTTCTCCCTTCTCATGAGCAAGAGGATTTCTTGATTG 1727

Db 1621 TACATATCAGGGGTTCTGTTCTCCCTTCTCATGAGCAAGAGGATTTCTTGATTG 1680

Qy 1728 CAGTCTAAAATTAATATTCAGTATGTTTGGCAAGCATGGAGTACCTTCAACACCTG 1787

Db 1681 CAGTCTAAAATTAATATTCAGTATGTTTGGCAAGCATGGAGTACCTTCAACACCTG 1740

Qy 1788 ACACAGGCAATTAATATTCAGTATGTTTGAACAGTCAATATTTCTTCTATGAGGATGG 1847

QY 1128 GATAATGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGCTCTAGTGGTGAATAA 1187
DB 1081 GATATGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGCTCTAGTGGTGAATAA 1140
QY 1188 GATGAGCAGACATGTTTGTATGTGGCTTATGAAAAAGGCGATGATGCCATTGTCACTC 1247
DB 1141 GATGAGCAGACATGTTTGTATGTGGCTTATGAAAAAGGCGATGATGCCATTGTCACTC 1200
QY 1248 CTGAAGCATTTATAAGACACACAGATGAATTCGCCCTGTAAAGATATTTCTCAGCCTGA 1307
DB 1201 CTGAAGCATTTATAAGACACACAGATGAATTCGCCCTGTAAAGATATTTCTCAGCCTGA 1260
QY 1308 GGAGATGGCTCCATGTGTCTGTTCATACCCCTTGGGGAAGATTAAAAAGCATGACAAA 1367
DB 1261 GGAGATGGCTCCATGTGTCTGTTCATACCCCTTGGGGAAGATTAAAAAGCATGACAAA 1320
QY 1368 GAGAAGCAGATATTTCTCTCTTAAGAGCTGGATTCCTTCAATTTCAATTTCTCAGCTC 1427
DB 1321 GAGAAGCAGATATTTCTCTCTTAAGAGCTGGATTCCTTCAATTTCAATTTCTCAGCTC 1380
QY 1428 TCAGAAATTCAGTTCCATGATATTTGGCTCAGGTTCTTTGGGAAAGTATATAAGGA 1487
DB 1381 TCAGAAATTCAGTTCCATGATATTTGGCTCAGGTTCTTTGGGAAAGTATATAAGGA 1440
QY 1488 CGATGCAGAAATAAATPAGTGGCTATAAAACGTTATCGAGCCCAATACCTACTGCTCCAAG 1547
DB 1441 CGATGCAGAAATAAATPAGTGGCTATAAAACGTTATCGAGCCCAATACCTACTGCTCCAAG 1500
QY 1548 TCAGATGATATGTTTTCGCGAGAGGTCCATTTCTGCGCAGCTCAATCATCCCTGC 1607
DB 1501 TCAGATGATATGTTTTCGCGAGAGGTCCATTTCTGCGCAGCTCAATCATCCCTGC 1560
QY 1608 GTAATTCAGTTTGGGTGCTGTGTTGAATGATCCAGCAGGTTTCCCATTTGCTCACTCAA 1667
DB 1561 GTAATTCAGTTTGGGTGCTGTGTTGAATGATCCAGCAGGTTTCCCATTTGCTCACTCAA 1620
QY 1668 TACATATCAGGGGTCTCTGTTCTCTCTCTCATGAGCAGAGGATTTCTGATTTG 1727
DB 1621 TACATATCAGGGGTCTCTGTTCTCTCTCTCATGAGCAGAGGATTTCTGATTTG 1680
QY 1728 CAGTCTAAATTAATTTATGTCAGTAGATGTTGCCAAAGGATGAGTACCTTCAACAACCTG 1787
DB 1681 CAGTCTAAATTAATTTATGTCAGTAGATGTTGCCAAAGGATGAGTACCTTCAACAACCTG 1740
QY 1788 ACACAGCAATATPACATCGTGAATGAAACATGTCACAATATTTCTCTATCAGGATGG 1847
DB 1741 ACACAGCAATATPACATCGTGAATGAAACATGTCACAATATTTCTCTATCAGGATGG 1800
QY 1848 CATGCTGTGTGGCAGATTTGGAGATCAAGATTTCTACAGTCTCTGATGAGACACAC 1907
DB 1801 CATGCTGTGTGGCAGATTTGGAGATCAAGATTTCTACAGTCTCTGATGAGACACAC 1860
QY 1908 ATGCAAAAACAACTGGGAACCTCCGTTGGATGGCTCCTGAGGTGTTACGCGAGTGCAC 1967
DB 1861 ATGCAAAAACAACTGGGAACCTCCGTTGGATGGCTCCTGAGGTGTTACGCGAGTGCAC 1920
QY 1968 CGGTACACATCAAGCAGATGTTCTCAGCTATGCTCTGTGTCTGTGGAAATTTCTCACT 2027
DB 1921 CGGTACACATCAAGCAGATGTTCTCAGCTATGCTCTGTGTCTGTGGAAATTTCTCACT 1980
QY 2028 GGCGAAATTCATTCGCTCATCTCAAGCAGCGGCTCGGCGAGCAGACATGCTTACACAC 2087
DB 1981 GGCGAAATTCATTCGCTCATCTCAAGCAGCGGCTCGGCGAGCAGACATGCTTACACAC 2040
QY 2088 CACATCAGACTCCCATTTGGCTATTCATTTCCAAAGCCCATATCACTCTCTGTGATAGA 2147
DB 2041 CACATCAGACTCCCATTTGGCTATTCATTTCCAAAGCCCATATCACTCTCTGTGATAGA 2100
QY 2148 GGGTGGAAACCATGCTCTGAGGAGGACCGCAATTTCTGAACTTCTCATGAAGTTAGAA 2207
DB 2101 GGGTGGAAACCATGCTCTGAGGAGGACCGCAATTTCTGAACTTCTCATGAAGTTAGAA 2160

QY 2208 GAGTGTCTCTGCAACATTTAGCTGATGTTCTCTGTCATCAAGTAAACAGCAGTGGTCTCTC 2267
DB 2161 GAGTGTCTCTGCAACATTTAGCTGATGTTCTCTGTCATCAAGTAAACAGCAGTGGTCTCTC 2220
QY 2268 TCACCTTCTTCTTCTTCTGATTTGCTGCTGTAACCCGGGAGGACCTGGCCGGAGTCAATGTG 2327
DB 2221 TCACCTTCTTCTTCTTCTGATTTGCTGCTGTAACCCGGGAGGACCTGGCCGGAGTCAATGTG 2280
QY 2328 GCAGCATTAAGAAGTCTGTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2387
DB 2281 GCAGCATTAAGAAGTCTGTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT 2340
QY 2388 TTGTCCCAAAAGTCTGGCAATATTTCTCTCAAGTCTCTTTTGGAGGAGATGAAAAA 2447
DB 2341 TTGTCCCAAAAGTCTGGCAATATTTCTCTCAAGTCTCTTTTGGAGGAGATGAAAAA 2400
QY 2448 AGTCTTCAATACACACCCATTTGACAAATATGCTATGCTATGCTATGCTATGCTATGCTCAATG 2507
DB 2401 AGTCTTCAATACACACCCATTTGACAAATATGCTATGCTATGCTATGCTATGCTCAATG 2460
QY 2508 CATTTTCATTTCTGCGAAATAGTAGCAGCTTTTGAGGACAGCAGC 2552
DB 2461 CATTTTCATTTCTGCGAAATAGTAGCAGCTTTTGAGGACAGCAGC 2505

RESULT 10
BD269561LOCUS BD269561 3026 bp DNA linear PAT 17-JUL-2003
DEFINITION Novel CARK protein and nucleic acid molecules and uses therefor.

ACCESSION

BD269561.1 GI:33079329

VERSION

JP 2002535962-A/3.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

Rattus norvegicus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 3026)

AUTHORS

Raju,J.

TITLE

Novel CARK protein and nucleic acid molecules and uses therefor

JOURNAL

Patent: JP 2002535962-A 3 29-OCT-2002;
MILLENNIUM PHARMACEUTICALS INC

COMMENT

OS Rattus norvegicus (rat)
PN JP 2002535962-A/3
PD 29-OCT-2002
PF 10-DEC-1999 JP 2000586772
PR 11-DEC-1998 US 60/111938,14-APR-1999 US 09/291839 PI
JEVASEELAN RAJU
PC C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/

PC 10 C1201/02.

PC C12Q1/68,G01N33/15,G01N33/50,G01N33/53,G01N33/68//A61K38/45,

PC A61K39/395,

PC A61K45/00,A61P9/04,A61P9/10,A61P35/00,A61P43/00,C12N15/00, PC

C12N5/00,

PC A61K37/52

CC Novel CARK protein and nucleic acid molecules and uses CC

therefor

FEATURES

source

FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..3026
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"

ORIGIN

Query Match 66.4%; Score 2010; DB 6; Length 3026;
Best Local Similarity 80.6%; Pred. No. 0;
Matches 2425; Conservative 0; Mismatches 570; Indels 13; Gaps 6;

QY 16 CGGCCCTGGAGAGGAAAGAACTTATAATAATGGGAATTTATAAATCTAGACCAACC 75

DB 29 CAGCACAGGAGAAAGAAAGAACTTCTTTAAAATGGGAATTTACAAATCCAGCACAC 88

Qy	76	AAACTTGTACTGATGAATGGAAGAAAAAGCTCAGTGAATCATATGTATTCACAATAGAAA	135
Db	89	AGACTGTGTTCTGATGAATGGAAGAAAGAACTTAGTGAATCTTACGCTATTATCATAGAAA	148
Qy	136	GAATTAGAGATGAACCTGCAGATCAAGGAAAAAGAACTGACAGAACTTAAGGAATATATTG	195
Db	149	GGCTGGAGGATAACTGCAGATCAAGAAAAATGAAATTTCAAGAACTAAAGGACACATCTTTG	208
Qy	196	GCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTTACCGCACTGAAATGGGTGTCTC	255
Db	209	GCTCTGATGAAGCCTTCAGTGAAGTCAATTTAAATTTACCGCAGAGCGTGGCTGTCCC	268
Qy	256	TACTTCAATTTATGTTGCAATTTGTGGAGGCAAGAAATCAATATTCGAATCTTTATGTTGA	315
Db	269	TGCTACACCTCTGCTGTCTGTGGCGGCAACAAGTCACATATCCGTGCCCTTATGTTAA	328
Qy	316	AGGGCTCGGCCATCTCGACTGACAGAAATGGATTTACAGCCTTGCATTTAGCAGTTT	375
Db	329	AGGGCTCGGTCATTCAGACTGACAGAAATGGGTTTCAAGCTCTGCACCTGGCCGTTT	388
Qy	376	ACAAGGATAATGAGAAATTGATCACTTCTCTGCTTCACAGTGGAGCTGATATACAGCAGG	435
Db	389	ACAAGGACAGCCCGAACTTATCACTTCACTGTTGCACAGCGAGCAGATGTCAGCAAG	448
Qy	436	TTGGATACGGTGGCTCACTGCCCTCCATATTCGCTACATAGCTGCCACCTAGAGCTG	495
Db	449	TGGGATACGGTGGCTCAAGCCCTCCACATAGCTCAATAGCTGGAACCCAGAGGCTG	508
Qy	496	CTGATGTGCTGTTGCAACATGAGAGCTAATGTCATATTTCAAGATGCAGTTTTTTTCACTC	555
Db	509	CAGAGTGTCTCAACAATGGGCCCAATGTGAATGTTCAAGATSCGTCTCTTCCACC	568
Qy	556	CATTGCAATTTGACGCTACTATGGAATGAACAGGTAACTCGCTTCTTTTGAATTTG	615
Db	569	CAGTGCACATTTGACGCTTACTATGGGACGAGCAGGTAAACAGTGTCTCTTTTGAAGTTG	628
Qy	616	GTGCTGATGTAATGTAAGTGTGCAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAA	675
Db	629	GTGCTGATGTAATGTAAGCGTGAGTTGGGAGCAGGCTCTGCACCTGCGCTCTGCNA	688
Qy	676	AAGGATCTTTGAATTTGCAAAATCTTTGATGGAAGAGCGCAAAAGCAATGTGAATG	735
Db	689	AGGGCTCTTTCAACATTTGTGAATCTCTGCTAGAGAGGAGCAAAAGCAGATGTGAACG	748
Qy	736	CTCAGATAAATGAAGACCATGTCCCATCTCAATTTCTGTTCTGATTTGGACACCATGATA	795
Db	749	CTCAGACAAATGAAGACCACTGCTCTGCACTTCTGTCTCGATTTGGACACCAACATA	808
Qy	796	TAGTTAAGTATCTGCTGCAAAAGTGAATTTGGAAGTTCAACCTCATGTGTGTAATCTATG	855
Db	809	TAGTGAGCTACCTGCTCCAGAGTCACTTAGAGTCCAGCCTCACGTCAATTAACATCTATG	868
Qy	856	GAGATACCCCTTACACCTGGCGTCTACAAATGGCAAAATTTGAAGTTGCCAAGGAATCA	915
Db	869	GTGACACTCCTTTGCAACCTGGCATGTACAAATGGAAATTTTGAAGTTGCCAAGGAAATG	928
Qy	916	TCCAAATATCAGGAACAGAAAGTCTGACTAAGGAAACATCTTTCAGTGAACACAGCTTTTC	975
Db	929	TCCAGGTAAACAGGAACCTGAAGTCTGACTAAGGAAACATCTTTCAGCGNAGCAGCTTTTC	988
Qy	976	ATAGTCTTGTACTTATGGCAAGACATTTGACTGTCAAAATTTCTTCTTGATCAGAAATG	1035
Db	989	ACAGTCTTGTACTTATGGCAAGAACATTCAGCTGGTCAAAATTTCTTCTTGATCAGAAATG	1048
Qy	1036	TCATAAACATCAACCAACCAAGAGATGGCACACTGATTTACACTCTGCTGCTTACC	1095
Db	1049	CTGTGAACATTAACCAACCAAGAGATGGCACAAGATTTGACCTCTGCTTGTCTACC	1108
Qy	1096	ACGGTCACATTCGCTTCTGTTTCAAGTCTTCTTCTTGGATAATGGAGCTGATATGAATCTAGTGG	1155
Db	1109	ACGGCCATATCGCTTCTGTTTCAAGTCTTCTTCTTGAATATGGTCAGATATGAATCTTGTGG	1168

1156	QY	CTTGTGATCCGACGAGTCTAGTGGTGA AAAAGATGACGACAGCATGTTTGTGATGTGGGCTT	1211
1169	Db	CTTGTGATCCGACGAGTCTAGTGGTGA AAAAGATGACGACAGCATGTTTGTGATGTGGGCTT	1228
1216	QY	ATGAAAAGGCGATGATGCCATTGTCCACTCTCGTGAAGCATTTATAAGAGACACACAAGATG	1275
1229	Db	ACGAAAAGGACATGATGCCATTGTTCACCTCTGAGCNCATACAGAGACCCGAGGAG	1288
1276	QY	AATTGCCCTGTAATGAATATTCTCAGCCTGGAGGAGATGGCTCTTATGTGTCTGTTCCAT	1335
1289	Db	AGCTGCCATTGTAACGAATATTCCAGCCTGGAGGAGATGGCTCTTATGTGTCTGTTCCCT	1348
1336	QY	CACCCCTTGGGAGAGATTTAAAAGCATGACAAAAGAGGACAGATATTCTCTCCCTAAAGAG	1395
1349	Db	CCCCCTTGGGACAGATTTAAAAGCATGACAAAAGAGGAGCATGTTCTCTCTCCGAGGG	1408
1396	QY	CTGGAATGCCCTTCCACATTTCCATCTTTCCAGCTCTCAGAAATTTGAGTTTCCATGAGATTATG	1455
1409	Db	CTGAACCTACCTCCCGCTTCCATCTCCAACTCTCCGAAATCGAGTTTCCACAGAGATTATCG	1466
1456	QY	GCTCAGGTTCTTTTGGGAAAGTATATAAAGGACGATGACAGAAATAAATAAGTGGCTATAA	1515
1469	Db	GCTCGGGTTCCTTTGGGAAAGTCTATAAAGGGCGATGCAAAATAAATAAGTGGCGATCA	1528
1516	QY	AACGTTATCGAGCCAAATACCTACTGTCTCCAAGTCAGATGTGGATATGTTTGGCCGAGAGG	1575
1529	Db	AACGATACCGAGCCAAACCTACTGTCTCCAAGTCAGACGTGGATATGTTTGGCCGAGAGG	1588
1576	QY	TGTCATCTCTGCGACGCTCAATCATCTCCCTGGGTAAATTCAGTTTGTGGGTGCTTGCTTGA	1635
1589	Db	TGTCATCTCTGCGACGCTCAACACCTCTGCGTGGTTCAGTTTGTGGGTGCTTGCTTGG	1648
1636	QY	ATGATCCCGACGATTTGGCAATTTGTCACCTCAATACATATCAGGGGTTCCTGTGTTCTCC	1695
1649	Db	ATGACCCCGACGATTTGGCAATTTGTCACCTCAATACATATCAGGGGTTCCTGTGTTCTCC	1708
1696	QY	TCCTTTCATGACGAGAAGAGGATTTTGATTTGCACTCTAAATTAATTTATGACGTAGATG	1755
1709	Db	TGCTTTCATGAACAGAGAGAAATTTTGATTTGCACTCTAAATTAATTTATGACGTAGAG	1768
1756	QY	TTGCCAAAAGCATGGAGTACCTTCAAACTGACAGACGCCAATTATACATCTGACCTTGA	1815
1769	Db	TTGCCAAGGGCATGGAGTACCTGACAGCTTGACCCAGCCAATCATACACCGGACCTGA	1828
1816	QY	ACAGTCACAATATTTCTTCTATGAGGATGGGCACTGCTGTGGTGAGAGATTTTGGAGAA	1875
1829	Db	ACAGCCACAATATTTCTGCTCTATGAGGATGGCCATGCTGTGGTGGCAGATTTTGGAGAA	1888
1876	QY	CAAGATTTCTACAGTCTCTGGATGAAGAACAATGACAAAACAACTGGGAACCTCCGTT	1935
1889	Db	CAAGATTTCTGCACTCCCTGGATGAAGACAACATGAACAAAGCAGCAGGGAACCTGCGCT	1948
1936	QY	GGATGGCTCTGAGGTGTTCCAGCAGTGCACTCCGTACACCATCAAGACAGATGTCTTCA	1995
1949	Db	GGATGGCCCTCTGAGTGTCTACAGTGCACGAGATACACCATCAAGGCTGATGTCTTCA	2008
1996	QY	GCTATGCTGTGTGTGTGGGAAATTTCTCATCTGGCGAAATTTCCATTCGCTCATCTCAAGC	2055
2009	Db	GTTACTCCCTGTGTCTGTGGGAGTCTCTCACTGGAGAAATTTCCATTCGCTCATCTCAAGC	2068
2056	QY	CAGCGGCTGGGACGACAGATGGCTTACCCACCATCAGACCTCCCATTTGCTATTCCA	2115
2069	Db	CAGCGGCTGCAGACGACAGATATGGGTATCACCAATCAGACGCCCATTCGCGTATTCCA	2128
2116	QY	TTCCCAAGGCCATATCATCTCTGATACGAGGGTGGAAACGATGCTCTGGAAGGAAGAC	2175
2129	Db	TCCCAAGGCCATCTCATCTCCCTGATACGGGCTTGGAAATGATGCTCTCTGGAAGGACGAC	2188
2176	QY	CCGAATTTCTGAAGTTGTATGAAGTTAGAAGTGTCTCTGCAACATTTGAGCTGATGT	2235
2189	Db	CAGAGTTCTCTGAAGTCGTTAGCAAACTGGAGAGTGGCTATGCAATGTGGAGCTCATGT	2248
2236	QY	CTCCTGCATCAAGTAAACAGCAGTGGGTTCTCTCTCACCTTCTCTTCTCTGATGTCCTGG	2295

2249	Db	CTCCAGCATCAAGTAAACAGCAGTGGCTCTCTGTCACTCTCTCTCTCCGATGGCTGC	2308
2296	Qy	TGAACGGGAGGACACTGGCCGAGTCATGTGCGCAGCATTRAGAAGTCGTTTCAATTTGG	2355
2309	Db	TGAGCCGGGAGGGCTGGCCGAGCAGCGTGCAGCCCTTACGAGCCGCTTTTGATTTGG	2368
2356	Qy	AATATGCTCTAAATGCAAGGTCCTATGCTGCTTTGTCCCAAGTGTGGACAATATTCCT	2415
2369	Db	AGTATGCCCTAATATGCAAGGTCCTATGCTGGGTGGTCCCAAGTGTGGAAACACTCTTA	2428
2416	Qy	CTCAAGGTCGTCTTTGGAGGAGATGAAAGAAGTCTTCAATACACACCCATTTGCAAAAT	2475
2429	Db	ATCCGGCCCTGTCTTTGGAGGAGATGAATAGGAGCACCCAGATATTCAACTGTGTGACAAAT	2488
2476	Qy	ATGGCTATGTATCCGATCCCATGAGCTCAATGATTTTCATTTCTTCCGCCGAATAGTAGCA	2535
2489	Db	ACGGCTATGTGTCGATCCCATGAGCTGACGACCTTCACCTCCCGCCAAAGACGACAGCA	2548
2536	Qy	GCCTTGAGGACAGCAGCTGACAGCATTGGCGGTATACCTAAGGAGAGTGTTTTCCCGGAA	2595
2549	Db	ACTTTGAGGACAGCACTGACAG-GTCTGGCATACACCTTAGGGGGGTCTCCCCATCAGG	2607
2596	Qy	CTGACAGCAACGATTTCCAAACCACGGCAAGGTGCTTCCAACTATAACATTTACTCTCAA	2655
2608	Db	CTGACAGCAGTGATTTTACCCATGGCAGGCTTGCTTCCCAATTATAACGGCCTGCCCTCTG	2667
2656	Qy	AGGTCCTCCTTAAATTTGGGCTTGTGTTTTACTTGTGCTATTATTAATTTCCCACTATTACGAGG	2715
2668	Db	AGGT-TTCTTCAATCGTCTGTGTTATCTTAAAGCTCGTTTAAATTTCCCTTCTACAGACAG	2726
2716	Qy	CTTTGGATTTGCTCCTAAAGGAATAATATGCAAAAGAACCAAGACAGAATGTATTAAGA	2775
2727	Db	GCCTTGACTCATGCC--AAGCCTGAAGTGTCAAAAGACGAGATACAGANTGTGCATGAGGA	2784
2776	Qy	ATTGTTTTTAATTTTGTAAATTAATAAATAATTTAGATCGTTACTTTGGAAATGGAGCCTTA	2835
2785	Db	ATTGTTCTTAGTTTGATATTTAAAGCCCTTAATTGCTCGGGCTGGGTTCAAA---TC	2840
2836	Qy	AGTCTGTGTGGACAGATAATAATATGTTTTCTCGGGCTGAATATATGATAGACTTGTGTT	2895
2841	Db	TGTGTAGATAGCTGGGTTGACCCCTTATGTATTTGTAGACCAAACTGTGTGGGCTTGTGT	2900
2896	Qy	TGACAG-CTATGGGTATTATTTCTTAGAACATTGTCATTTTCTTTTCTCATTTATGTTACT	2954
2901	Db	TGAGGTTCTCCTGTGGGTTCTTTAAAAACAACCTGGCTGATTTATCTCTGTGTGCTTTT	2960
2955	Qy	TCTAGTGTTTCACTCTGTGATTAAGAATCTTTTGGTGAAATAGAAAAAATAAAAAAAA	3014
2961	Db	----GTTGTATTCTCTGTATTAAGTCTCTTCGGTGATCTAGAAAAAATAAAAAAAA	3016
3015	Qy	AGGGCGGC	3022
3017	Db	GGGGCGCC	3024

```

RESULT 11
AR269113
LOCUS          3026 bp          DNA          linear          PAT 10-APR-2003
DEFINITION     Sequence 7 from patent US 6500654.
ACCESSION      AR269113
VERSION         AR269113.1  GI:296999956
KEYWORDS
SOURCE          Unknown.
SOURCE          Unknown.
ORGANISM       Unclassified.
REFERENCE       1 (bases 1 to 3026)
AUTHORS        Raju,J.
TITLE          CARK protein and nucleic acid molecules and uses therefor
JOURNAL         Patent: US 6500654-A 7 31-DEC-2002;
FEATURES        Location/Qualifiers
                    1..3026
                    /organism="unknown"
source

```


SOURCE	ORGANISM	Unknown.
REFERENCE	Unclassified.	
AUTHORS	1 (bases 1 to 3026)	
TITLE	Raju, J.	
JOURNAL	CARK protein and nucleic acid molecules and uses therefor	
FEATURES	Patent: US 6660490-A 7 09-DEC-2003;	
source	Location/Qualifiers	
	1..3026	
	/organism="unknown"	
	/mol_type="genomic DNA"	
ORIGIN		
	Query Match	66.4%; Score 2010; DB 6; Length 3026;
	Best Local Similarity	80.6%; Pred. No. 0;
	Matches 2425; Conservative	0; Mismatches 570; Indels 13; Gaps 6;
QY	16	CGGCCTCGAGAAAGGAGAAACTTATAATAATGGGAAATTATAAATCTAGACCAACC 75
Db	29	CAGCACAGGAGAAAGCAAGACTTCTTTAAATGGGAAATTACAAATCCAGACCACAC 88
QY	76	AAACTTGCTACTGATGAATGGAGAGAAAAGCTCAGTGAATCATATGTATTCACATAGAAA 135
Db	89	AGACTTGCTTGATGATGGAGAGAAAGTTAGTGAATCTTACGCTATATCATAGAAA 148
QY	136	GATTAGAAGATGACCTGCAGATCAAGGAGAAAAGAACTGACAGAACTAAGGAATATTTG 195
Db	149	GGCTGGAGGATAA CCTGCAGATCAAGAGAAATGAATTTCAAGAACTAAGGCACATCTTG 208
QY	196	GCTCTGATGAGGCTTCAGTAAAGTCAATTTAAATACCGCACTGAAATGGCTGTCTC 255
Db	209	GCTCTGATGAGGCTTCAGTGAAGTCAATTTAAATACCGCACTGAGGCTGTCTC 268
QY	256	TACTTCATTTATGTTGCAATTTGGAGGCAAGAAATCACATATCGAACTCTTATGTTGA 315
Db	269	TGCTACACTCTGCTGTGCTGTGGCGCAACAAGTCACATATCGGTGCGCTTAATGTTAA 328
QY	316	AGGGCTCGGCCATCTCGACTGACAGAAATGGATTTACAGCCTTGCATTTACAGCTTT 375
Db	329	AAGGGCTCGTCCATCCAGACTGACAGAAATGGGTTTCCAGCTTCGACCTTGGCCGCTT 388
QY	376	ACAAGGATAATGCAGAAATTGATCATCTTCTCTGCTTCACAGTGGAGCTGATATACAGCAGG 435
Db	389	ACAAGGACAGCCCGAACTTATCACTTCACTGTGTCACAGCGAGCAGATGTTACAGCAAG 448
QY	436	TTGGATAGGGTGGCCTCACTGCCCTCCATATTGCTACATAGCTGGCCACCTAGAGGCTG 495
Db	449	TGGGATACGGTGGCCTCACTGAGCCCTCCACATAGCTCAATAGCTGGACCCACAGGCGTG 508
QY	496	CTGATGTGCTGTTGCAACATGGAGCTAAATGCTCAATATTCAGATGACGTTTTTTTCACTC 555
Db	509	CAGAAATGCTGCTCAACATGGGSCCAATGTGAATGTTCAAGATGCGCTTCTTCAACC 568
QY	556	CATTGCAATTTGGACGCTGCTATGGACATGAACAGTGAATCTGCCCTCTTTTGAATTTG 615
Db	569	CAGTGCACATTTGACGCTTACTATGGGCACAGCAGTAA CCAAGTGTCTTTTGAAGTTTG 628
QY	616	GTGCTGATGTAATGTTAAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAA 675
Db	629	GTGCTGATGTTCAATGTTAAAGCGTGAAGTTGGGACAGGCTCTGACCTGGCCCTTCGAA 688
QY	676	AAGGATCTTGAATTTGGCAAACTCTTGATGGAAGAGGAGCAAGACAGATGTGAATG 735
Db	689	AGGGCTTCTTCAAATTTGAAACTCTCTGGTAGAAGAGGAGCAAGACAGATGTGAACG 748
QY	736	CTCAAGATAATGAAGACCATGTGCCACTCCATTTCTGTTCTCGATTTGGACACCATGATA 795
Db	749	CTCAGGACAAATGAAGACCATGTCCTCTCTGCAATTTCTGATTTGGACACCAACATA 808
QY	796	TAGTTAAGTATCTCGTGGAAAGTATTTGGAGTTCAACTCATGTTGTTAATCTATG 855
Db	809	TAGTACAGTACTCTCTCCAGATGACTTTAGAGTCCAGCTTCAGCTCATTAACATCTATG 868

Qy	856	GAGATACCCCTTACACCTGGCATGCTACAAATGGCAAAATTTGAAGTTGGCAAGAAATCA	915
Db	869	GTGACATCTCCTTTGGCACTGGCATGCTACAAATGGAAATTTTGAAGTTGGCAAGAAATG	928
Qy	916	TCCAATATCAGGAACAGAAAGTCGCTAAGAGAAACATCTTCAGTGAACAGCTTTTC	975
Db	929	TCAGGTAAACAGAACTGAAAGTCTGACTAGGAAACATCTTCAGCGACAGCTTTTC	988
Qy	976	ATAGTGTCTTACCTATGGCAAGCATTGACCTAGTCAAAATTTCTTTGTATCAGAATG	1035
Db	989	ACAGTGTCTTACTATGGCAAGCAATTGACCTGGTCAAATTTCTTTGTATCAGAATG	1048
Qy	1036	TCATARAATCAACCAACCAAGGAGGAGTGGGCACACTGGATTACACTCTGCTTGCTACC	1095
Db	1049	CTGTGAACATTAACCAACGAGGAGAGATGGGCACACAGGATTCGCTTCGTGTACC	1108
Qy	1096	ACGGTCACATTCGCCCTGGTTTCAGTTCCTTACTGATTAATGGAGCTGATATGAATCTAGTGG	1155
Db	1109	ACGCCATATCGCCCTGGTTTCAGTTCCTACTGATTAATGGTCAGATATGAATCTTGTG	1168
Qy	1156	CTTGTGATCCACAGAGGTCTAGTGGTGA AAAAGATGAGCAGACATGTTTGATGTGGCTT	1215
Db	1169	CTTGTGATCCACAGAGGTCTAGTGGTGA AAAAGATGAGCAGACATGTTTGATGTGGCTT	1228
Qy	1216	ATGAAAAGGGCATGATGCCATTGTACACTCTCTGAAGCATTATAAGAGACCAACAAGATG	1275
Db	1229	ACGGAAGAGCATGATGCCATTGTTACACTCTCTGAAGCATTACAGAGACCCCGAGAG	1288
Qy	1276	AATTGCCCTGTAAATGAATATCTCAGCCTGGAGGAGATGGCTCCTATGTGTCTGTTCAT	1335
Db	1289	AGCTGCCATGTAAAGCAATATTCACAGCCTGGAGGAGATGGCTCCTATGTGTCTGTTCCT	1348
Qy	1336	CACCTTGGGGAAGATTAAAGCATCACAAAAGAGAAAGCAGATATTCCTCCTCAAGAG	1395
Db	1349	CCCCCTTGGCAAGATTAAAGCATGACAAAAGAGAAAGCAGATGTTCTCTCTGAGGG	1408
Qy	1396	CTGGATGTCCTTCAATTTCCATCTTCAGCTCTCAGAAATTGAGTTCCATGAGATTATG	1455
Db	1409	CTGAACTACCCCTCCCGCTCCATCTCCAACTCTCCGAAATTCGAGTTCCACGAGATTATCG	1468
Qy	1456	GCTCAGGTCCTTTGGGAAGTATATAAGAGCGATGCAGAAATAAATAGTGGCTATAA	1515
Db	1469	GCTCGGGTTCCTTTGGGAAGTCTATAAGAGGGGATGCAGAAATAAATAGTGGCGATCA	1528
Qy	1516	AAGCTTATCAGAGCAATACCTACTGCTCCAAGTCAGATGTGGATATGTTTGGCCGAGAGG	1575
Db	1529	AACGATACCGAGCCACACTACTGCTCCAACTCAGCTGGATATGTTTGGCCGAGAGG	1588
Qy	1576	TGTCGATCTCTGCCAGCTCAATCATCCTCGGTAAATTCAGTTTGGGGTGCTTGCTTGA	1635
Db	1589	TGTCGATCTCTGCCAGCTCAACACCCCTGGGTGTTTGTGTTTGGGGTGCTTGCTTGG	1648
Qy	1636	ATGATCCGAGCAGTTTGCCATTGTCACATACATATCAGGGGGTCTCTCTGTTCTCCC	1695
Db	1649	ATGACCCGAGTCAGTTTGCCATTGTCACCTAGTACATTCAGGAGGCTCCCTGTTCTCCC	1708
Qy	1696	TCCTTCATGAGCAGAGAGAGATTCCTTGTATTTGAGTCTCAAAATTAATATTCAGTATG	1755
Db	1709	TGCTTCATGAACAGAGAGAAATTCCTGACTGTCAGTCTAAATTAATTCATTCGCGTAGACG	1768
Qy	1756	TTGCCAAAGCATGGAGTACCTTCACAACTGACACAGCAATTAACATTCGTGACTTGA	1815
Db	1769	TTGCCAAAGGCATGGAGTACCTTCGACAGCTTGACCCAGCCAAATCATACCCGCACTGA	1828
Qy	1816	ACAGTCACAATATTCCTCTATAGGATGGGCATGCTGTGTGGCAGATTTTGGAGAAAT	1875
Db	1829	ACAGCCACAATAATTCCTGCTCTATGAGATGGCCATGCTGTGTGGCAGATTTTGGAGAAAT	1888
Qy	1876	CRAGATTTCTACAGTCTCTCGGATGAGACAAATGACRAAACTGCGGACCTCCGTT	1935
Db	1889	CAAGATTTCTGAGTCCCTCGGATGAGACAAATGACRAAACTGCGGACCTCCGCT	1948
Qy	1936	GGATGGCTCTCGAGGTGTTTCAACGCACTGCACTCGGTACACCAATCAAGACAGATGCTTCA	1995

Db	1949	GGATGCCCTCTGAGGTGTTACAGTGCAGAGATACCAATCAAGCGTGAATGCTTCA	2008
Qy	1996	GCTATGCTCTGCTGTGCGGAAATTTCTCACTGCGGAAATTTCCATTCGCTCATCTCAAGC	2055
Db	2009	GTTACTCCTGTGCTGTGCGGCTCTCACTGCGGAAATTTCCATTCGCTCATCTCAAGC	2068
Qy	2056	CAGCGCTGCGGAGCAGACATATGGCTTACACACATCAGACCTCCCATTTGGCTATCCA	2115
Db	2069	CAGCGCTGCGGAGCAGACATATGGCTTACACACATCAGACCTCCCATTTGGCTATCCA	2128
Qy	2116	TTCCCAAGCCATATCATCTCTGCTGATACGAGGCTGAACGATGCTTGAAGGAGAC	2175
Db	2129	TCCCAAGCCATATCATCTCTGCTGATACGAGGCTGAACGATGCTTGAAGGAGAC	2188
Qy	2176	CCGAATTTTCTGAAGTCTCATGAAGTTGAAGAGTGTCTCTGCAACATTTGAGCTGATGT	2235
Db	2189	CAGAGTTCTCTGAAGTCTTGAACACTGGAGGAGTGCCTATGCAATGTGGAGCTCATGT	2248
Qy	2236	CTCCTGCATCAGTAACAGCAGTGGCTCTCTCACCTTCTTCTTCTGATTGCTCTG	2295
Db	2249	CTCCAGCATCAGTAACAGCAGTGGCTCTCTCACCTTCTTCTTCTGATTGCTCTG	2308
Qy	2296	TGAACCGGGGAGGAGCTCGCGGAGTCTATGTGGCAGCATTAAGAAGTCTGTTTCAATTTG	2355
Db	2309	TGAGCGGGGAGGAGCTCGCGGAGGAGTCTATGTGGCAGCATTAAGAAGTCTGTTTCAATTTG	2368
Qy	2356	AATATGCTCTAATCAAGTCTCTATGCTGCTTTGTCCTCAAGTCTGGACATATTCCT	2415
Db	2369	AGTATGCTCTAATCAAGTCTCTATGCTGCTTTGTCCTCAAGTCTGGACATATTCCT	2428
Qy	2416	CTCAAGTCTGCTTTGAGGAGATGAAGAAGTCTTCAATACACACACCATTCGACAAAT	2475
Db	2429	ATCCGGGCTGCTTTGAGGAGATGAAGAAGTCTTCAATACACACACCATTCGACAAAT	2488
Qy	2476	ATGGCTATGATTCGATCCCATAGCTCAATGCAATTTTCAATTTGTCGCGAATATGACGA	2535
Db	2489	ACGGCTATGATTCGATCCCATAGCTCAATGCAATTTTCAATTTGTCGCGAATATGACGA	2548
Qy	2536	GCTTTGAGCAGCAGCTGACAGCATTCGGGCTTACCTAAGGAGAGTCTTTTCCCGGAA	2595
Db	2549	ACTTTGAGCAGCAGCTGACAGCATTCGGGCTTACCTAAGGAGAGTCTTTTCCCGGAA	2607
Qy	2596	CTGACAGCAACGATTTCCACACGCGAGCTGGCTTCCAACTATAAATTTTACTCTCAA	2655
Db	2608	CTGACAGCAGTATTTTACCATGCGAGCTTGCTTCCAAATATAAAGCGCTTCCCTCTG	2667
Qy	2656	AGGTCTCTTAATTTGGGCTGTTTACTTGTCTTATTTAATTTCCCATATTAAGCAGG	2715
Db	2668	AGGT-ATCTTCAATTCGCTTGTCTTATTTAAGTCTGTTTAAATTTCCCTTCTACAGCAG	2726
Qy	2716	CTTTGATTTGCTTGAAGGATATATATGCAAAAGACCAAGACAGCAATCTATATGAAGA	2775
Db	2727	GCTTTGACTCATGCC--AGCCTGAAGTGTCAAAGAGCAGATACAGAAATGTCATGAGGA	2784
Qy	2776	ATTGTTTTTAATTTGTTAAATTAATAAATTTAGATCGTTACTTGGAAATGGAGCCTA	2835
Db	2785	ATTGTTCTTAGTTGATATTTAAAGCCCTTAATTTGCTGGGCTGGGGTTCAAA---TC	2840
Qy	2836	AGTCTGTGTGACAGATAATTAATTTTCTTCTGGGCTGAATATATGATGATTTGTTT	2895
Db	2841	TGTTAGATAGTGGTGGTACCCCTTATGTTTGTAGACCAACCTGTTGGGCTTGTGTT	2900
Qy	2896	TGACAG-CTATGGGTTTATTTCTTAGAACAATTTGTTTCTTTTCTCATTTGTTACT	2954
Db	2901	TGAGGCTCTCTGTTGGGTTTCTTAAACCAAGCTGGCTGATTTATCTCTGTTGCTTTT	2960
Qy	2955	TCTAGTTTCACTCTGTGATTAAGATTTCTTTGTTGAATATGAAAAAATAAATAAATAA	3014
Db	2961	---GTTGTTACTTCTGTGATTAAGTCTCTTCGGTGTATCTAGAAAAAATAAATAAATAA	3016
Qy	3015	AGGCGGCGC 3022	

Db

3017

GGGCGGCC 3024

RESULT 13

AY303692

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AY303692

Rattus norvegicus cardiac ankyrin repeat kinase mRNA, complete cds.

2982 bp

mRNA

linear

23-JUN-2003

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus

1 (bases 1 to 2982)

Jeyaseelan,R.

Rat cardiac ankyrin repeat kinase (rCARK)

Unpublished

2 (bases 1 to 2982)

Jeyaseelan,R.

Direct Submission

Submitted (20-MAY-2003) Cardiovascular Biology, Millennium Pharmaceuticals Inc., 75 Sidney Street, Cambridge, MA 02139, USA

Pharmaceuticals Inc.

Location/Qualifiers

1. 2992

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

44..2551

/note="CARK; protein kinase"

/codon_start=1

/product="cardiac ankyrin repeat kinase"

/protein_id="AAP72031.1"

/db_xref="GI:32165610"

tranelations="MGNYKSRPTQTCSDWKVKVSEVAIIERLEDNLQIKENEFQELRHIFSGDEAFSEVSNLYRTERGLSLHLCCVCGGNKSHIRALMLKGRPSRLTRNGVPAHLAVIKDPSPLIITSLHSGADVOQVGYGLTALHIAIAGHPEAAEVLLOHGANTNVQDAVFTPLHIAIYHEQVTSVLLKFGADVNVSQVGVDRPLHLASAKKFFNVLKLVESKADNAQVQNDHVPFLHFCRSFHHNIVSYLLQSDLEVQSHVHINIDYDTPLHLACVNGFEVAKETVQVTESTLTKENIFSETAFHAGTCKNIDVLLKQSLININHRGRDGHGTGHSACVGHGRLVQFLDNGADNMLVACDPSRSGDEKQTCMLMAYEKHDAIVTLAKHYRPOBELPCNEYSOPGGDGVSVSPGLKIKMTKEKADVLLKRAELSRFHQLUSEIEFELHIGSGSFGVKYGRCKNIVAIAKVRANTYCSKSDVMFCREVSILCQLNHPCVQVFGACLDPPSFAIVQYISGGSLSFLHHEQKRLIDLOSXLIIADVAKGMEYLHSLTQPIIHRDUNSHNILLIEDGHAVADPFGESFLASLDLDDNMTKQPGNLRWMAPEVFTQCTRYTIIKADVFSYSLCMLLELTGTEIPFAHLKPAADMAVHHIRPPIGYSIPKPIISLLIRGNWACPEGRPEFSEVSKLESCLCNLMSPASSNSSLSPSSSDCLLSRGGPGRSHVAALRSRFELEYALNARSYAGWSQSVGTHSNPGLSLEEMNRSTQSTVDKYGVSDPMSLTHLSRQDDSNFEDSN"

ORIGIN

Query Match

Best Local Similarity

Matches 2404; Conservative

0; Mismatches 566; Indels 13; Gaps 6;

65.8%; Score 1991.4; DB 10; Length 2982;

80.6%; Pred. No. 0;

16 CGGCCCTGGAGAGGAGAACTTATAATAATGGAAATTAATAATCTAGACCAACCC 75

12 CAGCAGGAGGAAAGCAAGACTTCTTTAAATGGGAATTAACAATCCAGCAACAC 71

76 AAATCTGTACTGATGAATGGAAGAAAAAGTCAAGTGAATCATATGTTATCACAATAGAA 135

72 AGACTTGTCTGATGAATGGAAGAAAGTCTTAGTGAATCTTACGCTATTATCATAGAAA 131

136 GATTAGAGATGACTCCAGTCAAGAAAAGACTGACAGAACTAAAGAAATATATTTG 195

132 GGCTGGAGGATAAAGCTGCAGATCAAGAAATGAATTTCAAGAACTAAGGCATCTTTG 191

196 GCTCTGATGAAGCCCTTCAGTAAAGTCAATTTAAATTAACCGACTGAAATGGGCTCTC 255

192 GCTCTGATGAAGCCCTTCAGTGAAGTCAAGTTAAATTAACCGCAGAGCGTGGCTCTCC 251

256 TACTTCATTTATGTTGATTTGTGGAGCAAGAAATCACATATTCGAACTCTTATTTGTA 315

Db 252 TGCTACACCTCTGCTGTCTCTGTGGCGCAACAGTCACATATCCGTGCGCTTATGTTAA 311
Qy 316 AAGGGCTCCGCCATCTCGACTGACAAAGAAATGATTTACAGCCTTGCAATTTAGCAGTTT 375
Db 312 AAGGGCTCCGTCCTCAGACTGACGAGAAATGGTTTCAGCTGTCACCTGGCCGTTT 371
Qy 376 ACAAGGAATATGACAGAAATGATCACTTCTCTGCTTACAGTGAGCTGATATACACAGG 435
Db 372 ACAGGACAGCCCGAACTTATCACTTCACTGTGACAGCGGAGAGATGTTACAGAA 431
Qy 436 TTGGATACGGTGGCTCACTGCGCTCCATATGCTACAATAGCTGBCCACTAGAGGCTG 495
Db 432 TGGGATACGGTGGCTCACAGCTTCCACATAGCTGCAATAGCTGACACCCAGAGGCTG 491
Qy 496 CTGATGTGCTGTGCAACATGAGAGCTAATGCTCAATATTCAGATGCGATTTTTCACCTC 555
Db 492 CAGAAGTGTGCTACAACTATGGGCCAATGTGAATGTTCAAGATGCGCTTCTTCACCC 551
Qy 556 CATTGCATATTTGACGCTACTATGGACATGAACAGGTAACTCGCCTTCTTTTGAATTTG 615
Db 552 CACTGCACATTTGACGCTACTATGGGCAAGCAGGTAAACAGTGTCTTTTGAAGTTG 611
Qy 616 GTGCTGATGTAAGTAAAGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGMAA 675
Db 612 GTGCTGATGTAAGTAAAGGCTGAAGTTGGGACAGGCTCTGCACTGCGCTCTGCAA 671
Qy 676 AAGGATTTCTTGAATATTTGCAAACTCTTTGATGGAAGAGGACCAAGCAGATGTGAATG 735
Db 672 AGGCTTTCTCAATTTGTGAACCTCTGCTGTAGAGAGGGACCAAGCAGATGTGAAG 731
Qy 736 CTCAGATATGAGACCATGTCCCACTCCATTTCTGTTCTCGATTTGACCAACCATGATA 795
Db 732 CTCAGGACAAATGAGACCAAGTCCCTCTGCACTTCTGTTCTCGATTTGACACCAACA 791
Qy 796 TAGTTAAGTATCTGCTGAAGTGAATTTGAAGTTCAGCTCATGTTGTTATATCTATG 855
Db 792 TAGTGACTACCTGCTCCAGATGACTTAGAGTCCAGCTTCAGTCAATTAACATCTATG 851
Qy 856 GAGATACCCCTTACACCTGCGATGTACAATGGCAAAATTTGAAGTTGCCAAGGAAATCA 915
Db 852 GTGACACTCTTTGCACCTGGCATGTACAAATGAAATTTGAAGTTGCCAAGGAAATG 911
Qy 916 TCCAAATACAGGACAGAAAGTCTGACTAGGAAACATCTTCAGTGAACAGCTTTTC 975
Db 912 TCCAGGTAAACAGGAACTGAAGTCTGACTAAGGAAACATCTTCAGCGAGACAGCTTTTC 971
Qy 976 ATAGTGTCTTGTACTATGGCAAGACATTTGACCTAGTCAAATTTCTTTGATCAGAATG 1035
Db 972 ACAGTGTCTTGTACTATGGCAAGACATTTGACCTAGTCAAATTTCTTTGATCAGAATG 1031
Qy 1036 TCATAAATCATCAACCAAGGAGGATGGGCACACTGGATTAACCTGCTGCTTACTAC 1095
Db 1032 CTGTGAACATTAACCAAGGAGAGATGGGCACACAGGATTTGCACTCTGCTTACTACC 1091
Qy 1096 ACAGTCACTTCCGCTGTTTCACTTCTTACTGATATGAGCTGATATGATCTAGTGG 1155
Db 1092 ACAGCCATATCCGCTGTTTCACTTCTTCTGATGATGTCAGATATGATCTGTCG 1151
Qy 1156 CTGTGATCCAGCAGTCTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTATGTTGGGCTT 1215
Db 1152 CTGTGATCCAGCAGTCTAGTGGTGAAGAAAGATGAGCAGACATGTTTGTATGTTGGCTT 1211
Qy 1216 ATGAAGAGGGCATGATGCTTGTACACTCTGAGCATTTATAGAGACCAAGATG 1275
Db 1212 ACAGAAAGGACATGATGCTTGTGATCTTCTGATGATGTCAGATATGATCTGTCG 1271
Qy 1276 AATTGCCCTGTAAATGATTTCTCAGCTGGAGGATGGCTCCATGCTGCTGTTCCAT 1335
Db 1272 AGCTGCCATGTACGAATATTTCCAGCTGGAGGATGGCTCCATGCTGCTGTTCTCT 1331
Qy 1336 CACCTTGGGGAAGATTAAGCATGACAAAGAGAGGACAGATTTCTCTCTTAAGAG 1395

Db 1332 CCCCTTGGGCAAGATTAAAGCATGACAAAGAGAGGAGCATGTTCTCTCTCTGAGGG 1391
Qy 1396 CTGAATTTGCTTCCATTTCCATCTTCAGCTCTCAGAAATTTAGTTTCCATGAGATTTTG 1455
Db 1392 CTGAATTTGCTTCCATTTCCATCTTCAGCTCTCAGAAATTTAGTTTCCATGAGATTTTG 1451
Qy 1456 GCTCAGGTTCTTTTGGGAAAGTATATAAAGGAGCATGACAGAAATAAATAGTGGCTATAA 1515
Db 1452 GCTCGGTTCTTTTGGGAAAGTCTATAAAGGCGCATGACAGAAATAAATAGTGGCGATCA 1511
Qy 1516 AACGTTATCAGGCAATPACCTACTGCTCCAAAGTCAGATGAGATATGTTTTCGAGAGG 1575
Db 1512 AACGATACCGAGCAACACCTACTGCTCCAAAGTCAGACGCTGATATGTTTTCGAGAGG 1571
Qy 1576 TGTCATTTCTGCGAGCTCAATCATCTCCGCTAATTCAGTTTGTGGTGTGCTTCTTGA 1635
Db 1572 TGTCATTTCTGCGAGCTCAACACCTGCGTGGTTTCTGTTTGTGGTGTGCTGCTGG 1631
Qy 1636 ATGATCCAGCCAGTTTGGCATTGTCACTCAATACATATCAGGGGTTCTCTGTTCTCC 1695
Db 1632 ATGACCCAGTCAGTTTGGCATTGTCACTCAGTACATTTTTCAGAGGCTCCCTGTTCTCC 1691
Qy 1696 TCCTTATGAGCAGAGAGGATTTCTTGAATTTGAGTCTTAAATTAATTTATGAGTAGATG 1755
Db 1692 TGCTTATGAACAGAGAGAAATTTCTGACTTGCAGTCTAAATTAATTAATTTGCGGTAGAG 1751
Qy 1756 TTGCCAAAGCATGAGTACTTTCACAACTGACACAGCAATTAATACATCTGACTTGA 1815
Db 1752 TTGCCAAGGCTGAGTACTTTCGACAGCTTGAACCGCAATTAATACACCGGACCTGA 1811
Qy 1816 ACAGTCAAAATTTCTTCTATGAGGATGGGATGCTGTGTGGCAGATTTTGGAGAT 1875
Db 1812 ACAGCCACAATTTCTGCTCTATGAGGATGGGATGCTGTGTGGCAGATTTTGGAGAT 1871
Qy 1876 CAGATTTCTACAGTCTCTGAGTGAAGCAACATGACAAACAACTGGGACCTCCGTT 1935
Db 1872 CAGATTTCTGAGTCCCTGGATGAAGCAACATGACAAAGCAGCAGGAACTTCGCGT 1931
Qy 1936 GAGTGGCTCCTGAGGTGTTTCAAGCTGCACTCGGTACACCATCAACAGCAGATGCTTCA 1995
Db 1932 GAGTGGCTCCTGAGGTGTTTCAAGCTGCAAGTACACCATCAAGGCTGATGCTTCA 1991
Qy 1996 GCTATGCTCTGTTGTTGGGAAATTTCTCAGTGGGAAATTTCCATTCGCTCATCTCAAGC 2055
Db 1992 GTTACTCCCTGCTGTGTGGGAGTCTCTCAGTGGGAAATTTCCATTCGCTCATCTCAAGC 2051
Qy 2056 CAGGGCTGGGACAGACATGGCTTACCAACATCAGACTCCCATTTGGCTATTCA 2115
Db 2052 CAGCGCTGCAGCAGACATATGGGCTATCAACATCAGACCGCCCATCGGCTATTCA 2111
Qy 2116 TTCCCAAGCCCATATCATCTCTGATACGAGGTTGGAACGCAATGCTCTGAAGGAGAC 2175
Db 2112 TCCCAAGCCCATATCATCTCTGATACGAGGTTGGAATGCATGCTCTGAAGGAGAC 2171
Qy 2176 CCAATTTCTGAAGTTGTCATGATAGAGAGTCTCTGCAACATTTGAGCTGATGT 2235
Db 2172 CAGATTTCTGAGTCTGTTAGCAAACTGAGGAGTGGCTATGCAATGTTGAGTCTATGT 2231
Qy 2236 CTCCTGCATCAAGTAAACAGCAGTGGTCTCTCTCAGCTTCTTCTTCTGATTTCCCTGG 2295
Db 2232 CTCAGCATCAAGTAAACAGCAGTGGTCTCTCTGACCTTCTCTTCTTCCGATTTCCCTGC 2291
Qy 2296 TGAACCGGGAGGACCTGGCGGAGTCTATGCGGAGCATTAAGAGTGTGTTTCGAATTTGG 2355
Db 2292 TGAGCGGGAGGAGGCTGGCGGAGCAGTGGGAGCTTACGGAGCGCTTTTGAATTTGG 2351
Qy 2356 AATATGCTCTAAATGCAAGTCTCTATGCTTGTCTTCCCAAAGTCTGGACAAATTTCT 2415
Db 2352 AGTATGCCCTAAATGCAAGTCTCTATGCTGGTGTGCCCAAAGTGTGGAACACACTCTA 2411
Qy 2416 CTCAGGCTGTCTTTTGGAGGAGATGAAAAGAGTCTTCAATACACACCCATTTGACAAAT 2475
Db 2412 ATCCGGGCTGTCTTTTGGAGGAGATGAATAGGAGCAGCAGTATTCAACTGTTGACAAAT 2471

[illegible]

```
QY 1008 CTAGTCAAAATTTCTTTGATCAGAATGTCTATAAATCAATCAACCAAGGAGGATGGG 1067
Db 961 CTGGTCAAAATTTCTTTGATCAGAATGTCTATAAATTAACACCGAGGAGATGGG 1020
QY 1068 CACACTGATTTACACTCTGCTTGTCTACACGGTACACATTTGCCCTGGTTAGTTCCTACTG 1127
Db 1021 CACACAGGATTTGCACTCTGCTTGTCTACACGGCCATATCCGCCCTGGTTAGTTCCTACTT 1080
QY 1128 GATAATGAGCTGATATGAATCTAGTGGCTTGTGATCCACGACAGGCTAGTGGTGAAAAA 1187
Db 1081 GATAATGAGCTGATATGAATCTTGTGCTTGTGATCCACGACAGGCTAGTGGTGAAAAA 1140
QY 1188 GATGACGACATGTTTGATGTGGCTTATGAAGGAGGATGATGCCAATGTGCACTC 1247
Db 1141 GATGACGACATGTTTGATGTGGCTTATGAGAAAGGACATGATGCCAATGTGTACACTC 1200
QY 1248 CTGAAGCATTAAGAGACCAAGATGAATTTGCCCTGTAATGAATTTCTCAGCCTGGA 1307
Db 1201 CTGAAGCATTAAGAGACCCAGAGAGGCTGCCATGTAAACGATATTTCCAGCCTGGA 1260
QY 1308 GGAGATGGCTCCTATGTGTCTGTTCCATCAACCTTTGGGGAAGATTAAAAAGCATGACAAA 1367
Db 1261 GGAGATGGCTCCTATGTGTCTGTTCCCTTCCCTTGGGCAAGATTAAAAAGCATGACAAA 1320
QY 1368 GAGAAAGCAGATATTTCTCCTTAAGAGCTGGATTGCTTCAATTTCCATCTTCAGCTC 1427
Db 1321 GAGAAAGCAGATGTTCTCTCTGAGGCTGAACACTACCTCCCGCTTCATCTTCCAACTC 1380
QY 1428 TCAGAAATGAGTTCCATGAGATTAATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487
Db 1381 TCCGAAATCGAGTTCCACGAGATTAATCGGCTCGGTTCTTTTGGGAAAGTATATAAGGG 1440
QY 1488 CGATGCGAATAAATAAGTGGCTATATAAGCTATGAGCCAAATACCTACTGCTCCAG 1547
Db 1441 CGATGCGAATAAATAAGTGGGATCAAAAGGATCAACGACCAACACTACTGCTCCAG 1500
QY 1548 TCAGATGTGATATGTTTGGCAGAGGCTTCCATCTCTGCGAGCTCAATCATCCCTGC 1607
Db 1501 TCAGACGTGATATGTTTGGCAGAGGCTTCCATCTCTGCGAGCTCAACCAACCTGCG 1560
QY 1608 GTAATTCAGTTTGGGTGCTTGTGAATGATCCGAGCAGTTGCCATTTGCACTCAA 1667
Db 1561 GTGGTTTCAGTTTGGGTGCTGCTGATGACCCAGTCAGTTTGCCATTTGCTACTCAG 1620
QY 1668 TACATATCAGGGGTTCTCTGTTCTCCCTCTTATGACGACAGAGGATTTCTGATTTG 1727
Db 1621 TACATTTACAGAGCTCCCTGTTCTCCCTGCTTATGAACAGAGAGAATTTCTGACTTG 1680
QY 1728 CAGTCTAAATTAATTAATGAGTATGTTTGCACAAAGGATGAGTACCTTCCACACCTG 1787
Db 1681 CAGTCTAAATTAATTAATGAGTATGTTTGCACAAAGGATGAGTACCTTCCACACCTG 1740
QY 1788 ACACAGCAATTAATATCTGACTTGAACAGTCACAATATTTCTTCTATGAGGATGGG 1847
Db 1741 ACCAGCAATTAATATCTACCCGACCTTGAACAGCACAATATTTCTGCTTATGAGATGGC 1800
QY 1848 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTTCTACAGTCTCTGATCAACACAAAC 1907
Db 1801 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTTCTGAGTCTCTGATGAAGCAAC 1860
QY 1908 ATGACAAACAAACCTGGAAACCTCGTTGATGGCTTCTGAGGTGTTTACGAGTGCAT 1967
Db 1861 ATGACAAAGCAGCAGGAAACCTCGCTGATGGCCCTTGAAGTGTTCACACAGTGCACG 1920
QY 1968 CGGTACACCATCAAGCAGATGCTTTCAGTATGCTGCTGCTGCTGGGAAATTTCTCACT 2027
Db 1921 AGATACACCATCAAGGCTGATGCTTTCAGTACTTCCCTGCTGCTGAGGCTCTCACT 1980
QY 2028 GGGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGCGAGCAGACATGCTTACCAAC 2087
Db 1981 GGAGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGCAGCAGAGATATGCGCTATCAC 2040
```

```
QY 2088 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTTCTGTGATACGA 2147
Db 2041 CACATCAGACCGCCCATCGCTATTCCATCCCAAGCCCATCTCATCTCCCTGTGTACGG 2100
QY 2148 GGTGGGAAGCATGCTCTCTGAAGAACCCGAAATTTTCTGAAGTTGTCAAGAGTTAGAA 2207
Db 2101 GGCTGGAAATGATGCTCTCTGAAGGACGACAGAGTTCTCTGAAGTCGTTAGCAAACTGGAG 2160
QY 2208 GAGTGTCTCTGCAACATTTGAGCTGATGTCTCTGCAATCAAGTAACAGCAGTGGTCTCTC 2267
Db 2161 GAGTGCCTATGCAATGTGGAGCTCATGTCTCCAGCATCAAGTAACAGCAGTGGCTCTCTG 2220
QY 2268 TCACCTTCTCTTCTTCTTCTGATTTGCTGTTGAAACCGGGAGGACCTGGCCGAGTCAATGTG 2327
Db 2221 TCACCTTCTCTTCTTCTTCTGATTTGCTGCTGAGCCGGGAGGCTGGCCGAGACCACTG 2280
QY 2328 GCACATTAAGAGTCTGTTTTCGAATTTGGAATATGCTCTAAATCGAAGTCTCTATGCTCT 2387
Db 2281 GCACCTTACGAGCCGCTTTTGAATTTGAGTGTGAGTATGCCCTAAATCGAAGTCTCTATGCTGG 2340
QY 2388 TTGTCCCAAAAGTGTGACAAATATTCCTCTCAAGTCTGTCTTTGGAGGAGATGAAAAA 2447
Db 2341 TGTCTCCAAAGTGTGGAACACACTCTAATCCGGGCTGTCTTTGGAGGAGATGAATAGG 2400
QY 2448 AGTCTTCAATACACACCCATTGACAAATATGCTGCTATGATCCGATCCCATGAGCTCAATG 2507
Db 2401 AGCACCCAGTATCAACTGTTGACAAATAGCGCTATGTGTCTGATCCCATGAGCCTGACG 2460
QY 2508 CATTTTCATTTCTTCCGAAATAGTAGCAGCTTTTGAGGACAGCAGC 2552
Db 2461 CACCTTCACCTCCGCAAGACGACAGCAACTTTTGAGGACAGCAAC 2505

RESULT 15
AR269114
LOCUS AR269114 2505 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 9 from patent US 6500654.
ACCESSION AR269114
VERSION AR269114.1 GI:29699957
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2505)
AUTHORS Raju, J.
TITLE CARK protein and nucleic acid molecules and uses therefor
JOURNAL Patent: US 6500654-A 9 31-DEC-2002;
FEATURES Location/Qualifiers
source 1..2505
/mol_type="genomic DNA"

ORIGIN
Query Match 62.6%; Score 1892.2; DB 6; Length 2505;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

QY 48 ATGGGAAATTTATAAATCTAGACCAACCCAAATCTGTACTGATGAATGGAGAAAAGTC 107
Db 1 ATGGGAAATTTACAAATCTCCAGACCAACACAGACTTGTCTGATGAATGGAGAAAAGTT 60
QY 108 AGTGAATCATATGTTTATCATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 167
Db 61 AGTGAATCTTACGCTATTATCATAGAAGGCTGGAGGATAACCTGCAGATCAAGAAAAAT 120
QY 168 GAATCAGACAGAACTAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAGTCAATTTA 227
Db 121 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCTTCAGTGAAGTCAATTTA 180
QY 228 AATTACCGCACTGAAATGGGCTGCTCTTACTTCAATTTATGTTGATTTGAGGAGCAAG 287
Db 181 AATTACCGCACAGAGCGTGGCTGTCCCTGCTACACCTTGTCTGTCTGTGGCGGCAC 240
```

QY	288	AAATCACATATTCGAACTCTTATGTTGAAAGGGCTCGGCCATCTCGACTGACAGAAAT	347	Db	1321	GAGAGGCAGATGTTCTCTCTCTGAGGGCTGAATACCTCCCGCTCCATCTCCAACTC	1380
Db	241	AAGTCACATATCCGTGCGCTTATGTTAAAGGGCTCGGTCCATCCAGACTGACGAGAAAT	300	QY	1428	TCAGAAATTCAGTTCATGAGATTATGCTCAGGTTCTTTTGGGAAAGATATATAAGGA	1487
QY	348	GGATTTACAGCTTCGATTTAGCAGTTTACAAAGGATAATGACAGAAATGATCTCTCTG	407	Db	1381	TCGAAATTCGAGTTCACGAGATTATCGGCTCGGGTTCCTTTGGGAAAGTCTATAAGGG	1440
Db	301	GGGTTTCAGCTCTGACCTGCGCTTTACAGGACAGCCGGAATTTATCATCTTCACTG	360	QY	1488	CGATGCAGAAATAAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGTCTCCAA	1547
QY	408	CTTCAGTGGAGCTGATATACAGCAGTTGGATACCGTGGCTCACTGCGCTCCATATT	467	Db	1441	CGATGCAGAAATAAAATAGTGGCGATCAACGATACCGAGCCAAACACCTACTGTCTCCAA	1500
Db	361	TTGCACAGCGGAGCAGATGTTTCAGCAAGTGGGATACGGTGGCTTCACAGCCCTCCACATA	420	QY	1548	TCAGATTCGATATGTTTTCGCGAGAGGTGTCATTTCTCTGCGAGCTCAATCATCCCTGC	1607
QY	468	GCTACAAATAGCTGGCCACTAGAGCTGCTGATGCTGTGTTGCAACATGAGAGCTAATGTC	527	Db	1501	TCAGACGTGATATGTTTTCGCGAGAGGTGTCATTTCTCTGCGAGCTCAACCAACCCCTGC	1560
Db	421	GCTGCAATAGCTGGACACCCAGAGGCTGCAGAGTGTCTGTACAAATGGGGCCAATGTG	480	QY	1608	GTAATTCAGTTTCGCGGTGCTGCTGTAATGATCCGAGCAGTTTGGCCATTTGCTCACTCAA	1667
QY	528	AAATATCAAGATGCAGTGTGTTTTCCTCCATTTGATATTCAGCGTACTATGACATGAA	587	Db	1561	GTGGTTTCAGTTTTCGCGGTGCTGCTGCGATGACCCAGTCAGTTTGGCCATTTGCTCACTG	1620
Db	481	AAATGTTCAAGATGCGCTCTCTTCCACCCACTGACATTCAGCGCTACTATGGGCACGAG	540	QY	1668	TACATATCAGGGGGTTCCTCTGTTTCTCCCTTCATGACAGAGAGGATTTCTGATTTG	1727
QY	588	CAGGTAACTGCGCTTCTTTTGAATTTGGTCTGATGTAATGTAAGTGTGAGTTGGA	647	Db	1621	TACATTTTCAGAGGCTCCCTGTTTCTCCCTGCTTCATGAAACAGAGAGATTTCTGACTG	1680
Db	541	CAGGTAACAGGTGCTCTTTTGAAGTTTGGTGTGATGTCATGTAATGTAAGCGGTGAGTTGG	600	QY	1728	CAGTCTAAATTAATTTGCGATGATGTTGCCAAGGCGATGAGTACCTTCACAACTG	1787
QY	648	GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTGAAATTTGCAAACTCTTGTATG	707	Db	1681	CAGTCTAAATTAATTTGCGGTAGACGTTGCCAAGGCGATGAGTACCTTCACAACTG	1740
Db	601	GACAGGCTCTGCACCTGCGCTCTGCAAGGCTTCTTCAACATTTGTAACCTCTCTGGTA	660	QY	1788	ACACAGCAATTTATACATCGTACTTGAACAGTCAACATATTTCTTCTCTATGAGGATGG	1847
QY	708	GAAAGAGGAGCAAAAGCAGATGTGAATGCTCAAGATTAATGAAGCAATGTCCTCACT	767	Db	1741	ACCAGCAATCATACACCGGACCTGGAACGCCACAAATATTTCTGCTCTATGAGATGGC	1800
Db	661	GAAAGAGGAGCAAAAGCAGATGTGAACGCTCAGGACCAATGAAGACCACTGCTCTGCA	720	QY	1848	CATCTCTGTGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGATGTAAGCAAC	1907
QY	768	TTCTGTTCTCGATTTGGACACATGATATGTTAGTATCTGCTGCAAGTGTATTTGGAA	827	Db	1801	CATCTCTGTGTGGCAGATTTTGGAGAAATCAAGATTTCTGAGTCCCTGGATGAAGCAAC	1860
Db	721	TTCTGTTCTCGATTTGGACACCAATATAGTAGTACCTGCTTCCAGAGTACTTAGAG	780	QY	1908	ATGACAAAACAACTGGGACCTCCGTTGGATGGCTCTGAGTGTGTTACGAGTGCCT	1967
QY	828	GTTCACCTCATGTTGTTTAAATATCTATGAGATACCCCTTTACACCTGGCATCTACAAT	887	Db	1861	ATGACAAAACAACTGGGACCTCCGTTGGATGGCTCTGAGTGTGTTACGAGTGCCT	1920
Db	781	GTCCAGCCTCAGCTCATTAACATCTATGTTGACACTCTTTTGCACCTGGCATGTACAAT	840	QY	1968	CGGTACACCATCAAGACAGATGTTTACGCTATGCTCTGTCTGTGGGAAATTTCTCACT	2027
QY	888	GGCAAAATTTGAAGTTGCCAAGGAATCATCAAAATATCAGGACAGAAATCTGCACTAAG	947	Db	1921	AGATACACCATCAAGGCTGATGTTTCAAGTTACTCCCTGTCTGTGGGAGTCTCTCACT	1980
Db	841	GGAAATTTTGAAGTTGCCAAGGAATTTGTCAGGTAAACAGAACTGAAAGTCTGACTAAG	900	QY	2028	GGCAAAATTTCAATTCGCTCATCTCAAGCAGCGGCTCGGCGACAGACATGCTTACCAAC	2087
QY	948	GAAACATCTTCAGTGAACACAGCTTTTCATAGTGTGTTGACCTATGCAAGACATTTGAC	1007	Db	1981	GGAGAAATTCATTCGCTCATCTCAAGCAGCGGCTCGGCGACAGACATGCTTACCAAC	2040
Db	901	GAAACATCTTCAGCAGACAGCTTTTCAAGTGTGTTGACCTATGCAAGACATTTGAC	960	QY	2088	CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCATCTCTCTGTGATACGA	2147
QY	1008	CTAGTCAAAATTTCTTTGATCAGAAATGTCTAAACATCAACCAAGGAGGATGGG	1067	Db	2041	CACATCAGACCGCCCATTCGCTATTTCCATCCCAAGCCCATCTCTCTGTGATACGA	2100
Db	961	CTGGTCAAAATTTCTTTGATCAGAAATGTCTGAAACATTAACCAAGGAGGATGGG	1020	QY	2148	GGGTGGAACCGCATGCTCTGAGGAGGAGCCGAAATTTCTGAAAGTGTCTGATGAGTAGAA	2207
QY	1068	CACATGGAATACATCTGCTGCTGCTACCAAGGCTCAGATTCGCTGCTTCAAGTCTTACTG	1127	Db	2101	GGGTGGAATGATGCTCTGAGGAGGAGCCAGAGTTCCTGAAAGTGTCTGAAAGTGTGAG	2160
Db	1021	CACACAGGATTCGATCTGCTTGTCTACCAAGGCTATCCGCTGCTTCAAGTCTTACTT	1080	QY	2208	GAGTGTCTCTCAACATTTGAGCTGATGTTCTCTGCTCATCAAGTAAACAGAGTGGGTCTCTC	2267
QY	1128	GATAATGAGCTGATATGATCTAGTGGCTGTGATCCCAAGGCTCTAGTGGTGAAGAA	1187	Db	2161	GAGTGGCTTATGATGAGGCTCATGTTCTCCAGCATCAAGTAAACAGAGTGGGTCTCTG	2220
Db	1081	GATAATGAGCTGATATGATCTAGTGGCTGTGATCCCAAGGCTCTAGTGGTGAAGAA	1140	QY	2268	TCACCTTCTCTTCTTCTGATGCTGGTGAACCGGGAGGAGCTGGCGGAGTCAATG	2327
QY	1188	GATGAGCAGACATGTTGATGAGGCTTATGAAAGGCGATGATGCCATTTGTCACACTC	1247	Db	2221	TCACCTTCTCTTCTTCTGATGCTGGTGAACCGGGAGGAGCTGGCGGAGGAGCAGT	2280
Db	1141	GATGAGCAGACATGTTGATGAGGCTTATGAAAGGCGATGATGCCATTTGTCACACTC	1200	QY	2328	GCAGCATTAAGAGCTGTTTTCGAAATGATGCTCTAAATCAAGGCTCTATGCTGCT	2387
QY	1248	CTGAACATTAAGAGACCAACAGATGAATTTGCCCTGTAAATGAATTTCTAGCTGGA	1307	Db	2281	GCAGCTTACCGAGCCGCTTTTGTGAGTATGCGCTTAATCAAGTCTCTATGCTGGG	2340
Db	1201	CTGAACATTAAGAGACCAACAGATGAATTTGCCCTGTAAATGAATTTCTAGCTGGA	1260	QY	2388	TTGTCCCAAGTGTGGAACAATATTCCTCTCAAGGCTGTCTTTTGGAGGAGATGAAGAA	2447
QY	1308	GAGATGGCTCCTATGTTGCTGTTCCATCACCCTTGGGAGATTAAGGATGACAAAA	1367	Db	2341	TGGTCCCAAGTGTGGAACAACACTCTAAATCCGGGCTGTCTTTGGAGGAGATGAATAG	2400
Db	1261	GGAGATGGCTCCTATGTTGCTGTTCTTCCCTTGGGCAAGATTAAGGATGACAAAA	1320	QY	2448	AGTCTTCAATACACCCCATTTGCAATATGGCTATGATATCCGATCCCATGAGCTCAATG	2507
QY	1368	GAGAGGCGAGATATTCCTCTCTAAGAGCTGGATGCTTCCATTTCCATCTTTCAGCTC	1427				

Db 2401 AGCACCCAGTATTCAACTGTGTGACAAATACGGCTATGTGTCTGATCCCATGAGCCTGACG 2460
QY 2508 CATTTTCATTCTTCCGGAATAGTAGCAGCTTTCAGGACAGCAGC 2552
Db 2461 CACCTTCACTCCCGCCAAGACGACAGCACTTTGAGGACAGCAAC 2505

Search completed: September 6, 2004, 05:03:12
Job time : 7790.57 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2004, 05:03:46 ; Search time 757.842 Seconds
(without alignments)
16430.147 Million cell updates/sec

Title: US-10-626-173-3

Perfect score: 2505

Sequence: 1 atgggaataataaattctag.....gcagtttgaggacagcgc 2505

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3267054 seqs, 2485319735 residues

Total number of hits satisfying chosen parameters: 6534108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2505	100.0	2505	9	US-09-947-199-3
2	2505	100.0	2505	17	US-10-626-173-3
3	2505	100.0	3025	9	US-09-947-199-1
4	2505	100.0	3025	17	US-10-626-173-1
5	1892.2	75.5	2505	9	US-09-947-199-9
6	1892.2	75.5	2505	17	US-10-626-173-9
7	1892.2	75.5	3026	9	US-09-947-199-7
8	1892.2	75.5	3026	17	US-10-626-173-7
9	427.4	17.1	616	9	US-09-833-381-1149
10	414.6	16.6	736	9	US-09-833-381-1150
11	413.2	16.5	740	9	US-09-833-381-1151
12	184.2	7.4	304	9	US-09-833-381-1147
13	109.6	4.4	1953	13	US-10-424-599-53404
14	107.4	4.3	3238	13	US-10-424-599-129156

15	100.2	4.0	2207	13	US-10-425-114-22955	Sequence 22955, A
16	100.2	4.0	2327	17	US-10-437-963-54440	Sequence 54440, A
17	98	3.9	1682	9	US-09-938-842A-1014	Sequence 1014, Ap
18	98	3.9	1682	11	US-09-938-842A-1014	Sequence 1014, Ap
19	97	3.9	1572	13	US-10-425-114-33402	Sequence 33402, A
20	91.4	3.6	1369	13	US-10-425-114-30542	Sequence 30542, A
21	91.4	3.6	1545	13	US-10-425-114-21691	Sequence 21691, A
22	91.4	3.6	1614	13	US-10-425-114-13674	Sequence 13674, A
23	91.4	3.6	3323	16	US-10-104-047-313	Sequence 313, App
24	89.4	3.6	972	13	US-10-425-114-6886	Sequence 6886, Ap
25	88.6	3.5	1146	17	US-10-437-963-59479	Sequence 59479, A
26	88.6	3.5	1461	13	US-10-425-114-4043	Sequence 4043, Ap
27	87.8	3.5	1219	13	US-10-424-599-130828	Sequence 130828, A
28	87.8	3.5	3020	17	US-10-437-963-65894	Sequence 65894, A
29	87.4	3.5	1299	9	US-09-908-711-12	Sequence 12, Appl
30	87.2	3.5	2919	17	US-10-437-963-70898	Sequence 70898, A
31	85.6	3.4	5338	13	US-10-342-887-63	Sequence 63, Appl
32	85.6	3.4	5338	13	US-10-172-118-63	Sequence 63, Appl
33	85	3.4	1942	13	US-10-425-114-28428	Sequence 28428, A
34	85	3.4	1997	13	US-10-425-114-30767	Sequence 30767, A
35	84.4	3.4	186	9	US-09-833-381-1148	Sequence 1148, Ap
36	84.2	3.4	5175	9	US-09-964-899-42	Sequence 42, Appl
37	84	3.4	799	13	US-10-425-114-30676	Sequence 30676, A
38	84	3.4	14770	10	US-09-873-367C-230	Sequence 230, App
39	84	3.4	14770	12	US-09-968-007A-124	Sequence 124, App
40	84	3.4	14770	12	US-09-968-007A-405	Sequence 405, App
41	84	3.4	14770	13	US-10-342-887-1848	Sequence 1848, Ap
42	84	3.4	14770	13	US-10-172-118-1848	Sequence 1848, Ap
43	83.6	3.3	2473	17	US-10-437-963-92781	Sequence 92781, A
44	83.4	3.3	1991	17	US-10-437-963-93604	Sequence 93604, A
45	82.6	3.3	4213	16	US-10-108-260A-794	Sequence 794, App

ALIGNMENTS

RESULT 1

US-09-947-199-3
; Sequence 3, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: WNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199-3

Query Match 100.0%; Score 2505; DB 9; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGGAATTTATTAATCTAGACCAACCCCAACTTCTACTGTATGATGAAGAAAAAGTC 60

Db 1 ATCGGAATTTATTAATCTAGACCAACCCCAACTTCTACTGTATGATGAAGAAAAAGTC 60

Qy 61 AGTGAATCATATGTTTATCATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAAA 120


```
QY 2281 GCAGCATTAAAGAGTCGTTTCGAATATGCTCTMAATGCAAGGTCCTATGCTGCT 2340
D 2281 GCAGCATTAAAGAGTCGTTTCGAATATGCTCTMAATGCAAGGTCCTATGCTGCT 2340
QY 2341 TTGTCCCAAGTGTGACAAATATTCCTCAAGGTCGTCTTTGAGGAGATGAAAGA 2400
D 2341 TTGTCCCAAGTGTGACAAATATTCCTCAAGGTCGTCTTTGAGGAGATGAAAGA 2400
QY 2401 AGTCTTCAATACACACCAATGACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460
D 2401 AGTCTTCAATACACACCAATGACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460
QY 2461 CATTTTCATTCTTGCCGAATATGAGCAGCTTTGAGACAGCAGC 2505
D 2461 CATTTTCATTCTTGCCGAATATGAGCAGCTTTGAGACAGCAGC 2505

RESULT 2
US-10-626-173-3
; Sequence 3, Application US/10626173
; Publication No. US20040110232A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/10/626,173
; FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: US/09/947,199A
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2505)
US-10-626-173-3

Query Match 100.0%; Score 2505; DB 17; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAATTAATACTAGACCAACCCAACTTGACTGATGGAGGAAAGTC 60
D 1 ATGGGAAATTAATACTAGACCAACCCAACTTGACTGATGGAGGAAAGTC 60
QY 61 AGTGAATCATATGTTATCAATAGAAAGATAGAGATGACCTGAGATCAAGGAAAA 120
D 61 AGTGAATCATATGTTATCAATAGAAAGATAGAGATGACCTGAGATCAAGGAAAA 120
QY 121 GAATGACAGAACTAAGGAATATTTGGCTCTGATGAAGCCTTCAATGAAGTCA 180
D 121 GAATGACAGAACTAAGGAATATTTGGCTCTGATGAAGCCTTCAATGAAGTCA 180
QY 181 AATTACCGCACTGAAATGGGCTCTCTACTTTTATGTTGCTATTTGCTGGAGCAAG 240
D 181 AATTACCGCACTGAAATGGGCTCTCTACTTTTATGTTGCTATTTGCTGGAGCAAG 240
QY 241 AATACATATTCGAATCTTTATGTTGAAAGGCTCCGCCATCTCGACTGACAGAAAT 300
D 241 AATACATATTCGAATCTTTATGTTGAAAGGCTCCGCCATCTCGACTGACAGAAAT 300
QY 301 GGATTTACGCTTGCATTTAGCAGTTTACAGGATTAATGCAAGATTTGATCACTTCTG 360
D 301 GGATTTACGCTTGCATTTAGCAGTTTACAGGATTAATGCAAGATTTGATCACTTCTG 360
```

```
Db 301 GGATTTACGCTTGCATTTAGCAGTTTACAGGATTAATGCAAGATTTGATCACTTCTG 360
QY 361 CTTACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 420
D 361 CTTACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCACTGCCCTCCATATT 420
QY 421 GCTCAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGCTAATGTC 480
D 421 GCTCAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGCTAATGTC 480
QY 481 AATATTCAGATGCAGTCTTTTCACTCCATGCAATTTGACAGCGTACTATGGACATGA 540
D 481 AATATTCAGATGCAGTCTTTTCACTCCATGCAATTTGACAGCGTACTATGGACATGA 540
QY 541 CAGTAACTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGAAGTTGA 600
D 541 CAGTAACTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGAAGTTGA 600
QY 601 GATAGACCCCTCACCTAGCATCTGCAAAAGGATTTTGAATTTGCAAACTCTTGATG 660
D 601 GATAGACCCCTCACCTAGCATCTGCAAAAGGATTTTGAATTTGCAAACTCTTGATG 660
QY 661 GAAGAAGGACAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCTCCAT 720
D 661 GAAGAAGGACAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCTCCAT 720
QY 721 TCTGTCTCGATTTGGACACCATGATATAGTAAATGCTGCTGCAAGATGTTGAA 780
D 721 TCTGTCTCGATTTGGACACCATGATATAGTAAATGCTGCTGCAAGATGTTGAA 780
QY 781 GTTCAACCTCATGTTGTAATATCTATGAGATACCCCTTACCTGCGCATGCTACAT 840
D 781 GTTCAACCTCATGTTGTAATATCTATGAGATACCCCTTACCTGCGCATGCTACAT 840
QY 841 GGCATAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTG 900
D 841 GGCATAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTG 900
QY 901 GAAACATCTTCAGTGAACACAGCTTTTATAGTGTGTTGACCTATGGCAAGACATG 960
D 901 GAAACATCTTCAGTGAACACAGCTTTTATAGTGTGTTGACCTATGGCAAGACATG 960
QY 961 CTAGTCAAAATTTCTCTGATCAGAAATGTCATAAATCATCAACCAAGGAGGATGG 1020
D 961 CTAGTCAAAATTTCTCTGATCAGAAATGTCATAAATCATCAACCAAGGAGGATGG 1020
QY 1021 CACTGGATTTACATCTGCTGCTACACGGTCAATTCGCTGCTGCTGCTGCTGCTGCT 1080
D 1021 CACTGGATTTACATCTGCTGCTACACGGTCAATTCGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 GATAATGGAGCTGATGATGATCTAGTGGCTTCTGATCCAGCAGCTTAGTGGTGA 1140
D 1081 GATAATGGAGCTGATGATGATCTAGTGGCTTCTGATCCAGCAGCTTAGTGGTGA 1140
QY 1141 GATGAGCAGATGTTGATGTTGGCTTATGAAAGGCGCATGATGCCATGTCACATC 1200
D 1141 GATGAGCAGATGTTGATGTTGGCTTATGAAAGGCGCATGATGCCATGTCACATC 1200
QY 1201 CTGAGCAGATTAAGAGACCAAGATGAAATTTGCCCTGTATGATTAATGATTTCT 1260
D 1201 CTGAGCAGATTAAGAGACCAAGATGAAATTTGCCCTGTATGATTAATGATTTCT 1260
QY 1261 GAGATGGCTCTTATGCTGCTGCTTCCATCAGCTTTGGGAAAGATTTAAAGCAT 1320
D 1261 GAGATGGCTCTTATGCTGCTTCCATCAGCTTTGGGAAAGATTTAAAGCAT 1320
QY 1321 GAGAGGCGAGATTTCTCTCTAAGAGCTGATGCTTCAATTTCCATTTCCATCTT 1380
D 1321 GAGAGGCGAGATTTCTCTCTAAGAGCTGATGCTTCAATTTCCATTTCCATCTT 1380
QY 1381 TCAGAAATTTGAGTTCATGAGATTTTGGCTCAGGTTCTTTTGGGAAAGATTA 1440
D 1381 TCAGAAATTTGAGTTCATGAGATTTTGGCTCAGGTTCTTTTGGGAAAGATTA 1440
```

QY	1441	CGATGCGAATAAATAAGTGGCTATAAAACGTTATTCGAGCCAATACCTACTGCTCCAAG	1500
Db	1441	CGATGCGAATAAATAAGTGGCTATAAAACGTTATTCGAGCCAATACCTACTGCTCCAAG	1500
QY	1501	TCAGATGTGGATATGTTTTCGCGAGAGGTGTCACATTCCTGCGCAGCTCAATCATCCTCTGC	1560
Db	1501	TCAGATGTGGATATGTTTTCGCGAGAGGTGTCACATTCCTGCGCAGCTCAATCATCCTCTGC	1560
QY	1561	GTAAATTCAGTTTGTGGGTGCTTGGTTGAATGATCCAGCCAGTTTGCCATTGTCACCTCAA	1620
Db	1561	GTAAATTCAGTTTGTGGGTGCTTGGTTGAATGATCCAGCCAGTTTGCCATTGTCACCTCAA	1620
QY	1621	TACATATCAGGGGGTTCCTGTGTTCTCCCTCTTCATGAGCAGAAGAGATTCTTGATTG	1680
Db	1621	TACATATCAGGGGGTTCCTGTGTTCTCCCTCTTCATGAGCAGAAGAGATTCTTGATTG	1680
QY	1681	CAGTCTAAATTAATTAATGCGAGTAGATGTTTGCCAAAGGCATGGAGTACCTTCACAACCTG	1740
Db	1681	CAGTCTAAATTAATTAATGCGAGTAGATGTTTGCCAAAGGCATGGAGTACCTTCACAACCTG	1740
QY	1741	ACACAGCCAAATTAACATCGTGACTCGAACAGTCACAATATTTCTCTATAGAGGATGGG	1800
Db	1741	ACACAGCCAAATTAACATCGTGACTCGAACAGTCACAATATTTCTCTATAGAGGATGGG	1800
QY	1801	CATGCTGTGGCGAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC	1860
Db	1801	CATGCTGTGGCGAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC	1860
QY	1861	ATGACAAAACAACTGGGAACTCCGTTGGATGGCTCCTGAGGTGTTCAACGAGTGCAC	1920
Db	1861	ATGACAAAACAACTGGGAACTCCGTTGGATGGCTCCTGAGGTGTTCAACGAGTGCAC	1920
QY	1921	CGGTACACCATCAAAAGCAGATGCTTCAGCTATGCTCTGTGTCGTGGGAAATTTCTCACT	1980
Db	1921	CGGTACACCATCAAAAGCAGATGCTTCAGCTATGCTCTGTGTCGTGGGAAATTTCTCACT	1980
QY	1981	GGCGAAATTCATATCGCTCATCTCAAGCCAGCGGCTCGCGCAGCAGACATGGCTTACAC	2040
Db	1981	GGCGAAATTCATATCGCTCATCTCAAGCCAGCGGCTCGCGCAGCAGACATGGCTTACAC	2040
QY	2041	CACATCAGACCTCCGATTTGGCTATTCATTCGACGCCATATCATCTCTGCTGATACGA	2100
Db	2041	CACATCAGACCTCCGATTTGGCTATTCATTCGACGCCATATCATCTCTGCTGATACGA	2100
QY	2101	GGTGGAAACGATGTCTGTAAGGAAGACCCGAAATTTCTGAAGTTGTCTGAAGTTAGAA	2160
Db	2101	GGTGGAAACGATGTCTGTAAGGAAGACCCGAAATTTCTGAAGTTGTCTGAAGTTAGAA	2160
QY	2161	GAGTGTCTCGAACATTTGAGTGTCTCTGCAATCAAGTAAACAGCAGTGGGTCTCTC	2220
Db	2161	GAGTGTCTCTGCAACATTTGAGTGTCTCTGCAATCAAGTAAACAGCAGTGGGTCTCTC	2220
QY	2221	TCACTTCTTCTTCTTCTGATTGCTGGTGAACCGGGGAGGACCTGGCCGAGTCATGTG	2280
Db	2221	TCACTTCTTCTTCTTCTGATTGCTGGTGAACCGGGGAGGACCTGGCCGAGTCATGTG	2280
QY	2281	GCAGATTAAAGAGTCGTTTCGAATTTGGAAATATGCTCTAAATGCAAGTCTCTATGCTGCT	2340
Db	2281	GCAGATTAAAGAGTCGTTTCGAATTTGGAAATATGCTCTAAATGCAAGTCTCTATGCTGCT	2340
QY	2341	TTGTGCCAAAGTGTGACAAATATTCCTCTCAAGGTCTGTCTTTGGAGGAGATGAAAGA	2400
Db	2341	TTGTGCCAAAGTGTGACAAATATTCCTCTCAAGGTCTGTCTTTGGAGGAGATGAAAGA	2400
QY	2401	AGTCTTCATATACACACCATTCGAAATATGGCTATGATCCGATCCCATGAGTCAATG	2460
Db	2401	AGTCTTCATATACACACCATTCGAAATATGGCTATGATCCGATCCCATGAGTCAATG	2460
QY	2461	CATTTTCATTTCTTGCCGAAATAGTAGCAGCTTTTGAGGACAGCAGC	2505
Db	2461	CATTTTCATTTCTTGCCGAAATAGTAGCAGCTTTTGAGGACAGCAGC	2505

```

RESULT 3
US-09-947-199-1
; Sequence 1, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju. Jeyaseelan
; TITLE OF INVENTION: NOVEL CARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48) ..(2552)
US-09-947-199-1

```

Query Match	100.0%	Score 2505	DB 9	Length 3025
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2505	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	ATGGGAATATATAACTCTAGACCAACCCAAACTTGTACTGATGAATGGAAGAAAAAGTC	60	
Db	48	ATGGGAATATATAACTCTAGACCAACCCAAACTTGTACTGATGAATGGAAGAAAAAGTC	107	
Qy	61	AGTGAATCATATGTTATCTCAATAGAAAGATTAGAGATGACCTCGACATCAAGGAAAA	120	
Db	108	AGTGAATCATATGTTATCTCAATAGAAAGATTAGAGATGACCTCGACATCAAGGAAAA	167	
Qy	121	GAACCTGACAGAACTAAGGAATATATTGGCTCTGATGGAAGCCTTCAGTAAAGTCAATTTA	180	
Db	168	GAACCTGACAGAACTAAGGAATATATTGGCTCTGATGGAAGCCTTCAGTAAAGTCAATTTA	227	
Qy	181	AATTACCCGACCTGAAATAGGGCTGCTCTACTTCTATTATGTTGATTTGTTGGAGGCAAG	240	
Db	228	AATTACCCGACCTGAAATAGGGCTGCTCTACTTCTATTATGTTGATTTGTTGGAGGCAAG	287	
Qy	241	AAATCACATATTGGAACCTTTATGTTGAAAGGCTCGCCCATCTCGACTGACAGAAAT	300	
Db	288	AAATCACATATTGGAACCTTTATGTTGAAAGGCTCGCCCATCTCGACTGACAGAAAT	347	
Qy	301	GGATTTACAGCCTTGCAATTTAGCAGTTTACAAGGATAATGCAGAAATTGATCACTTCTCTG	360	
Db	348	GGATTTACAGCCTTGCAATTTAGCAGTTTACAAGGATAATGCAGAAATTGATCACTTCTCTG	407	
Qy	361	CTTCAAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCTCACTGCCCTCGATATT	420	
Db	408	CTTCAAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCTCACTGCCCTCGATATT	467	
Qy	421	GCTACAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTTGCAACATGGAGCTAAATGTC	480	
Db	468	GCTACAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTTGCAACATGGAGCTAAATGTC	527	
Qy	481	AATATTCAAGATGCAGTTTTTTTTTCACTCCATTCGATTCGATTCGACCGTACTATGGACATGAA	540	
Db	528	AATATTCAAGATGCAGTTTTTTTTTCACTCCATTCGATTCGATTCGACCGTACTATGGACATGAA	587	
Qy	541	CAGGTAACCTCGCCTTCTTTTGGAAATTTGGTGCATGTAAATGTAAAGTGGTGAAGTTGGA	600	
Db	588	CAGGTAACCTCGCCTTCTTTTGGAAATTTGGTGCATGTAAATGTAAAGTGGTGAAGTTGGA	647	

QY 601 GATAGACCCCTCCACCTAGCATCTGCAAAAAGGATTTCTTGAATATTCGAAAACTCTTGATG 660
DB 648 GATAGACCCCTCCACCTAGCATCTGCAAAAAGGATTTCTTGAATATTCGAAAACTCTTGATG 707
QY 661 GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720
DB 708 GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 767
QY 721 TTCTGTCTCGATTTGAGCAACCATGATATAGTAAATCTGCTGCAAAAGTATTTGGAA 780
DB 768 TTCTGTCTCGATTTGAGCAACCATGATATAGTAAATCTGCTGCAAAAGTATTTGGAA 827
QY 781 GTTCAACCTCATGTGTTAAATATCTATGAGATACCCCTTACACCTGCGCATGCTACAT 840
DB 828 GTTCAACCTCATGTGTTAAATATCTATGAGATACCCCTTACACCTGCGCATGCTACAT 887
QY 841 GGCAAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAG 900
DB 888 GGCAAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAG 947
QY 901 GARAACATCTTCAGTGAACAGCTTTTCATAGTGGCTGTACCTATGGCAAGGACATTGAC 960
DB 948 GAACAACATCTTCAGTGAACAGCTTTTCATAGTGGCTGTACCTATGGCAAGGACATTGAC 1007
QY 961 CTAGTCAAAATTTCTTCTGTGATCAGAAATGTCATAAATCAACCAAGCAAGGATGGG 1020
DB 1008 CTAGTCAAAATTTCTTCTGTGATCAGAAATGTCATAAATCAACCAAGCAAGGATGGG 1067
QY 1021 CACACTGGATTAACACTCTGCTGTGATCCAGCGGTACATTCGCGCTGTTTCACTTCTACTG 1080
DB 1068 CACACTGGATTAACACTCTGCTGTGATCCAGCGGTACATTCGCGCTGTTTCACTTCTACTG 1127
QY 1081 GATAATGGAGCTGATAGTAATAGTGGCTGTGATCCAGCAGGTCTAGTGTGAAAA 1140
DB 1128 GATAATGGAGCTGATAGTAATAGTGGCTGTGATCCAGCAGGTCTAGTGTGAAAA 1187
QY 1141 GATGAGCAGACATGTTTGTGATGGCTTTATGAAAAAGGCGATGACCATTTGTCACTC 1200
DB 1188 GATGAGCAGACATGTTTGTGATGGCTTTATGAAAAAGGCGATGATGCCATTTGTCACTC 1247
QY 1201 CTGAGCAATTAAGAGACACAGATGATTTGCCCTGTAAATGATTTCTCAGCTCGA 1260
DB 1248 CTGAGCAATTAAGAGACACAGATGATTTGCCCTGTAAATGATTTCTCAGCTCGA 1307
QY 1261 GGAGATGGCTCCATATGTGCTGTTCATCACCCCTTGGGAAGATTAAGAGCATGACAAA 1320
DB 1308 GGAGATGGCTCCATATGTGCTGTTCATCACCCCTTGGGAAGATTAAGAGCATGACAAA 1367
QY 1321 GAGAAGCAGATATTTCTCTCTTAAGAGCTGATGCTTCCATTTCCATTTCCATTTCTCAGCTC 1380
DB 1368 GAGAAGCAGATATTTCTCTCTTAAGAGCTGATGCTTCCATTTCCATTTCCATTTCTCAGCTC 1427
QY 1381 TCAGAAATTCAGTTCATGAGATTTATGGCTCAGGTTCTTTTGGGAAGTATATATAAGGA 1440
DB 1428 TCAGAAATTCAGTTCATGAGATTTATGGCTCAGGTTCTTTTGGGAAGTATATATAAGGA 1487
QY 1441 CGATCGAGAAATAAATAGTGGCTATATAAAGCTTATCGAGCCAAATACCTACTGCTCCAAG 1500
DB 1488 CGATCGAGAAATAAATAGTGGCTATATAAAGCTTATCGAGCCAAATACCTACTGCTCCAAG 1547
QY 1501 TCAGATCTGATATGTTTTCGCGAGAGTGTCCATTTCTGCGCAGCTCAATCATCCCTGC 1560
DB 1548 TCAGATCTGATATGTTTTCGCGAGAGTGTCCATTTCTGCGCAGCTCAATCATCCCTGC 1607
QY 1561 GTAATTCAGTTTGGGTGCTTGTGTTGAATGATCCAGCAGTTTGGCCATTTGCTACTCAA 1620
DB 1608 GTAATTCAGTTTGGGTGCTTGTGTTGAATGATCCAGCAGTTTGGCCATTTGCTACTCAA 1667
QY 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGAGCAGAGGATTTCTGATTTG 1680
DB 1668 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGAGCAGAGGATTTCTGATTTG 1727
QY 1681 CAGTCTAAATTAATTAATTCGAGTAGATGTTTGGCAAGGATGGATGCTTCCAAACCTG 1740

DB 1728 CAGTCTAAATTAATTAATTCGAGTAGATGTTGCAAAAGGATGGAGTACCTTCCAAACCTG 1787
QY 1741 ACAGAGCCAAATTAATACATCGTGAATTTGAACAGTCAAAATTTCTTCTATGAGGATGG 1800
DB 1788 ACAGAGCCAAATTAATACATCGTGAATTTGAACAGTCAAAATTTCTTCTATGAGGATGG 1847
QY 1801 CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1860
DB 1848 CATGCTGTGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGCAAC 1907
QY 1861 ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTACGCAAGTGCATC 1920
DB 1908 ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTACGCAAGTGCATC 1967
QY 1921 CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTGCTGTGTGTGGAAATTTCTACT 1980
DB 1968 CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTGCTGTGTGTGGAAATTTCTACT 2027
QY 1981 GGGGAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGCAGCAGACATGCTTACCAC 2040
DB 2028 GGGGAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGCAGCAGACATGCTTACCAC 2087
QY 2041 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCATATCTCTGCTGATAGA 2100
DB 2088 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCATATCTCTGCTGATAGA 2147
QY 2101 GGGTGGAAACATGCTCTGAAAGAACCCGGAATTTCTGAAAGTTGTCATGAAGTAGNA 2160
DB 2148 GGGTGGAAACATGCTCTGAAAGAACCCGGAATTTCTGAAAGTTGTCATGAAGTAGNA 2207
QY 2161 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTGTCATCAAGTAAACAGCAGTGGGCTCTC 2220
DB 2208 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGTCATCAAGTAAACAGCAGTGGGCTCTC 2267
QY 2221 TCACCTTCTTCTTCTCTGATTTGCTGTTGAAACCGGGAGGACCTGCGCGAGTCAATG 2280
DB 2268 TCACCTTCTTCTTCTCTGATTTGCTGTTGAAACCGGGAGGACCTGCGCGAGTCAATG 2327
QY 2281 GCAGCATTAAGAAAGTCTGTTTGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCT 2340
DB 2328 GCAGCATTAAGAAAGTCTGTTTGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGCT 2387
QY 2341 TTGTCCAAAGTCTCGCAATATTTCTCTCAAGGTCCTCTTTTGGAGGAGATGAAAA 2400
DB 2388 TTGTCCAAAGTCTCGCAATATTTCTCTCAAGGTCCTCTTTTGGAGGAGATGAAAA 2447
QY 2401 AGCTTTCATACACACCCATTGACAAATATGCTATGCTATGCTATGCTATGCTCAATG 2460
DB 2448 AGCTTTCATACACACCCATTGACAAATATGCTATGCTATGCTATGCTATGCTCAATG 2507
QY 2461 CATTTTCATTTCTTGGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505
DB 2508 CATTTTCATTTCTTGGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2552

RESULT 4

US-10-626-173-1
; Sequence 1, Application US/10626173
; Publication No. US20040110232A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MN-068CP2
; CURRENT FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: US 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457

Db 1968 CGGTACACCAATCAAGAGAGATGCTTCAGCTATAGCTCTGTGTCTGTGGGAAATTTCTCACT 2027
Qy 1981 GGGCAAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGCGCAGCAGACATGGCTTACCAC 2040
Db 2028 GGGCAAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGCGCAGCAGACATGGCTTACCAC 2087
Qy 2041 CACATCAGACCTCCATTCGCTATTCATTCACAGCCCATATCATCTCTGCTGATACGA 2100
Db 2088 CACATCAGACCTCCATTCGCTATTCATTCACAGCCCATATCATCTCTGCTGATACGA 2147
Qy 2101 GGGTGGAAAGCATCTCTGAAGGAAGACCCGAATTTCTGAAAGTCTCATGAAGTTAGAA 2160
Db 2148 GGGTGGAAAGCATCTCTGAAGGAAGACCCGAATTTCTGAAAGTCTCATGAAGTTAGAA 2207
Qy 2161 GAGTGTCTGCAACATGAGCTGATCTCTCTGCAATCAAGTAAACAGCAGTGGGTCTCTC 2220
Db 2208 GAGTGTCTGCAACATGAGCTGATCTCTCTGCAATCAAGTAAACAGCAGTGGGTCTCTC 2267
Qy 2221 TCACCTCTCTCTCTCTCTGATTCGCTGTAACCGGGGAGCCTGGCGGAGTCAATG 2280
Db 2268 TCACCTCTCTCTCTCTGATTCGCTGTAACCGGGGAGCCTGGCGGAGTCAATG 2327
Qy 2281 GCAGCATTAAAGAGTCTGTTTCAAGTAAAGTATGCTCTAAATGCAAGTCTTATGCTGCT 2340
Db 2328 GCAGCATTAAAGAGTCTGTTTCAAGTAAAGTATGCTCTAAATGCAAGTCTTATGCTGCT 2387
Qy 2341 TTGTCCAAAGTCTGACAAATATCTCTCAAGGCTGTCTTTGGAGAGATGAAAGA 2400
Db 2388 TTGTCCAAAGTCTGACAAATATCTCTCAAGGCTGTCTTTGGAGAGATGAAAGA 2447
Qy 2401 AGTCTTCAATACACACCATGTAACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460
Db 2448 AGTCTTCAATACACACCATGTAACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2507
Qy 2461 CATTTTCATCTTCCGAAATAGTAGAGCTTTTGGAGACAGCAGC 2505
Db 2508 CATTTTCATCTTCCGAAATAGTAGAGCTTTTGGAGACAGCAGC 2552

RESULT 5

US-09-947-199-9
; Sequence 9, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL DARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MMI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; PRIOR FILING DATE: 2001-09-05
; PRIOR FILING DATE: 1998-12-11
; PRIOR FILING DATE: 1998-12-11
; PRIOR FILING DATE: 1999-04-14
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199-9

Query Match 75.5%; Score 1892.2; DB 9; Length 2505;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 1 ATGGGAATTAATCTAGCCCAACCAACCACTGTACTGATCAATGGAGAAAAGTC 60
|||||

Db 1 ATGGGGAATTTACAAATCCAGACCAAACACAGACTTGTCTGATGAATGGGAAGAAGTT 60
Qy 61 AGTGAATCATATGTTATCACAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAAA 120
Db 61 AGTGAATCTTACCTATATCAAGAAAGCTGGAGATAAAGCTGCAGATCAAGGAAT 120
Qy 121 GAACGTACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAAGTAAGTCAATTTA 180
Db 121 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCCTTCAAGTAAGTCAATTTA 180
Qy 181 AATTACCGCACTGAATAAGGCTGTCTTACTTCAATTTATGTTGCAATTTGAGAGCAAG 240
Db 181 AATTACCGCAAGAGCGTGGCTGTCTTACTTCACTCTGCTACACTCTGCTGTCTGTGGGCAAC 240
Qy 241 AATTCACATATTCGAATCTTATGTTGAAAGGGCTCCGCCATCTTCGACTGACAGAAAT 300
Db 241 AAGTCAATATCCGTGCCCCATGTTTAAAGGGCTCCGTCCATCCAGACTGACGAGAAAT 300
Qy 301 GGAATTTACAGCTTGCAATTTAGCAGTTTACAGGATAATGCAGAAATGATGATCACTTCTCTG 360
Db 301 GGGTTTCCAGCTCTGACCTGGCCGTTTACAGGACAGCCCGGAATCTTACATCTCACCTG 360
Qy 361 CTTACAGTGGAGCTGATATACAGCAGGTGGATACGGTGGCTCACTGCCCTCCATATT 420
Db 361 TTGCACAGCGAGCAGATGTTTCAAGAAAGTGGGATACGGTGGCTCACAGCCCTCCACATA 420
Qy 421 GCTACATAGCTGGCCACCTAGAGGCTGCTGATGCTGTGTTGCAACATGGAGCTAATGTC 480
Db 421 GCTGCAATAGCTGGACACCCAGAGGCTGCAGAGTGTCTCTACCAATGGGCGCAATGTC 480
Qy 481 AATATTCAAGATGCAAGTCTTCTTCACTCCATTCGATTCGATTCGAGCGTACTATGCAATGAA 540
Db 481 AATGTTCAAGATGCGCTCTCTTCACTCCATTCGATTCGAGCGTACTATGCGCACAG 540
Qy 541 CAGGTAACTCGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGA 600
Db 541 CAGGTAACTCGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGG 600
Qy 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATCTTGAATATTGCAAAATCTTTGATG 660
Db 601 GACAGGCTCTGCACCTGGGCTCTGCAAGGGGCTTCTTCAATGTTGAATCTCTGTTA 660
Qy 661 GAAGAAGGCAAGAGCAGATGTGAACGCTCAGGACAAATGAAGACCCACCTCTCTGCAC 720
Db 661 GAAGAAGGCAAGAGCAGATGTGAACGCTCAGGACAAATGAAGACCCACCTCTCTGCAC 720
Qy 721 TCTGTTCTCGATTTGGACACCATGATATAGTAAATGTAAGTATCTGCTGCAAGTCAATTTGAA 780
Db 721 TCTGTTCTCGATTTGGACACCAATATAGTAAATGTAAGTATCTGCTGCAAGTCAATTTGAA 780
Qy 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAT 840
Db 781 GTCCAGCTCATGTTAATATCTATGGTACATCTTGGTACATCTCTTGGACCTGGCATGCTACAT 840
Qy 841 GGCATAATTTGAAGTTCGAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
Db 841 GGAATAATTTGAAGTTCGAAGGAAATTTGTCAGGTAACAGGAACTGAAAGTCTGACTAAG 900
Qy 901 GAAACATCTTCAAGTAAACAGCTTTTCAAGTGTGTGCTTACCTATGCGCAAGAGCTTAC 960
Db 901 GAAACATCTTCAAGTAAACAGCTTTTCAAGTGTGTGCTTACCTATGCGCAAGAGCTTAC 960
Qy 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCATAAATCAACCAAGGAGGATGGG 1020
Db 961 CTGCTCAAAATTTCTTCTGATCAGAAATGTCGTGAACATTAACCAAGGAGAGATGGG 1020
Qy 1021 CACACTGGATTACATCTGCTTGTACAGGCTGATTCGGCTGGTTCAGTCTTACTCTG 1080
Db 1021 CACACAGGATTGACATCTGCTTGTACAGGCTGATTCGGCTGGTTCAGTCTTACTCTT 1080
Qy 1081 GATAATGGAGCTGATGATGATCTAGTGGGCTTGTGATCCAGCAGAGGCTTAGTGGTGA 1140
Db 1081 GATAATGGTGCAGATGATGATCTTGTGCTTGTGATCCAGCAGGCTTAGTGGTGA 1140

```
QY 1141 GATGAGCAGACATGTTTGATGTGGCTTATGAAGAGGCGATGATGCCATTGTCACTC 1200
Db 1141 GATGAGCAGACATGTTTGATGTGGCTTACGAGAGGACATGATGCCATTGTTCACATC 1200
QY 1201 CTGAAGCATTAAGAGACCAACAGATGAATGGCCCTGTATGAATATCTCAGCCTGGA 1260
Db 1201 CTGAAGCATTAAGAGACCCAGAGAGCTGCCATGTAAAGAAATATTCAGCCTGGA 1260
QY 1261 GGAGATGGCTCCTATGTGTCTGTTCCATCACCCTTGGGGAAGATTAAGAGCATGACAAA 1320
Db 1261 GGAGATGGCTCCTATGTGTCTGTTCCATCACCCTTGGGGAAGATTAAGAGCATGACAAA 1320
QY 1321 GAGAGGCGAGATATCTCTCTCTAGAGCTGGATGCTTCAATTTCCATCTCAGCTC 1380
Db 1321 GAGAGGCGAGATATCTCTCTCTAGAGCTGGATGCTTCAATTTCCATCTCAGCTC 1380
QY 1381 TCAGAAATGAGTTCATGAGATATTTGGCTCAGGTTCTTTGGGAAAGTATATAAGGA 1440
Db 1381 TCAGAAATGAGTTCATGAGATATTTGGCTCAGGTTCTTTGGGAAAGTATATAAGGA 1440
QY 1441 CGATGAGAAATAAATAGTGGCTATAAAGCTTATGAGCCAAATACCTACTGCTCAAG 1500
Db 1441 CGATGAGAAATAAATAGTGGCTATAAAGCTTATGAGCCAAATACCTACTGCTCAAG 1500
QY 1501 TCAGATGTGGATATGTTTGGCGAGAGGTGCCATTTCTGCCAGCTCAATCATCCTGTC 1560
Db 1501 TCAGATGTGGATATGTTTGGCGAGAGGTGCCATTTCTGCCAGCTCAATCATCCTGTC 1560
QY 1561 GTAATTCAGTTTTCGGGTCTGTTGTAATGATCCAGCCAGTTTGCAATGTCACCTCA 1620
Db 1561 GTGGTTTCAGTTTTCGGGTCTGTTGTAATGATCCAGCCAGTTTGCAATGTCACCTCA 1620
QY 1621 TACATATCAGGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
Db 1621 TACATATCAGGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
QY 1681 CAGTCTAAATTAATTTGTCAGTATGATTTGCCAAGGCGATGGAGTACCTTCAACCTG 1740
Db 1681 CAGTCTAAATTAATTTGTCAGTATGATTTGCCAAGGCGATGGAGTACCTTCAACCTG 1740
QY 1741 ACAGGCCAATATACATGCTGATTTGAAAGTACATATTTCTCTCTATGAGGATGG 1800
Db 1741 ACCAGGCCAATATACATGCTGATTTGAAAGTACATATTTCTCTCTATGAGGATGG 1800
QY 1801 CATGCTGTGTCGCGAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACA 1860
Db 1801 CATGCTGTGTCGCGAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACA 1860
QY 1861 ATGACAAAACAACTGGGAACTCTCGTTGGATGAGTCTCTGAGGTGTTTCAAGAGTGC 1920
Db 1861 ATGACAAAACAACTGGGAACTCTCGTTGGATGAGTCTCTGAGGTGTTTCAAGAGTGC 1920
QY 1921 CGGTACACCATCAAGCAGATGTTTCAAGTATGCTCTGCTCTGTCGGGAAATTCCTCA 1980
Db 1921 AGATACACCATCAAGCAGATGTTTCAAGTATGCTCTGCTCTGTCGGGAAATTCCTCA 1980
QY 1981 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGCTCGGCGAGCAGACATGGCTTACCAC 2040
Db 1981 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGCTCGGCGAGCAGACATGGCTTACCAC 2040
QY 2041 CACATCAGACCTCCATTTGGTATTTCCATTCGAGCCCATATCATCTCTGCTGATACGA 2100
Db 2041 CACATCAGACCTCCATTTGGTATTTCCATTCGAGCCCATATCATCTCTGCTGATACGA 2100
QY 2101 GGGTGGAGCGCATGTTCTGAGGAGAGACCCGAATTTTCTGAAGTTGTTCAGTAAGTAGAA 2160
Db 2101 GGGTGGAGCGCATGTTCTGAGGAGAGACCCGAATTTTCTGAAGTTGTTCAGTAAGTAGAA 2160
QY 2161 GAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGATCAAGTACAGCAGTGGGTCTCTC 2220
Db 2161 GAGTGTCTCTGCAACATTTGAGCTGATGTTCTCTGATCAAGTACAGCAGTGGGTCTCTC 2220
```

RESULT 6

```
US-10-626-173-9
; Sequence 9, Application US/10626173
; Publication No. US20040110232A1
```

GENERAL INFORMATION:

```
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MM1-068CP2
; CURRENT APPLICATION NUMBER: US/10/626,173
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: US/09/947,199A
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 60/111,938
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: US 09/458,457
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO: 9
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
```

US-10-626-173-9

```
Query Match 75.5%; Score 1892.2; DB 17; Length 2505;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;
```

```
QY 1 ATGGGAATTAATAATCTAGACCAACCCAAACCTTTGATGATGGAAGAAAAGTGC 60
Db 1 ATGGGAATTAATAATCTAGACCAACCCAAACCTTTGATGATGGAAGAAAAGTGC 60
QY 61 AGTGAATCATATGTTATCACAATAGAAAGATTAGAAGTGAAGTGCAGATCAAGGAAA 120
Db 61 AGTGAATCATATGTTATCACAATAGAAAGATTAGAAGTGAAGTGCAGATCAAGGAAA 120
QY 121 GAATGACAGCACTTAAGGATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTT 180
Db 121 GAATGACAGCACTTAAGGATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTT 180
QY 181 AATTACCGCACTGAAAATGGGCTGCTCTACTTTTATTTTGTGCAATTTGTGGAGGCAAG 240
Db 181 AATTACCGCACTGAAAATGGGCTGCTCTACTTTTATTTTGTGCAATTTGTGGAGGCAAG 240
QY 241 AATACATATTTGAACTCTTATGTTGAAGGCTCCGCCATCTCGACTGACAGAAAT 300
Db 241 AATACATATTTGAACTCTTATGTTGAAGGCTCCGCCATCTCGACTGACAGAAAT 300
```

QY 301 GGATTTACAGCCTTGCATTTAGCAGTTTACAAGATAATGCAGAAATGATCACTTCTCTG 360
Db 301 GGGTTTCAGCTGTCACCTGGCCGTTTTTACAAGACAGCCGGAATATATCACTTCACTG 360
QY 361 CTTACAGTGGAGCTGATATACAGCAGGTTGGATAGCGTGGCTCACTGCCCTCCATATT 420
Db 361 TTGCACAGCGGACAGATGTTACAGAGTGGATACGGTGGCTCACAGCCCTCCACATA 420
QY 421 GCTACAATAGTGGCCACCTAGAGGCTGCTGATGCTGTGTGCAACATGGAGCTAATGTC 480
Db 421 GCTGCAATAGTGGACACCCAGAGGCTGCAAGAGTGTGCTGCTACAACATGGGCCCATTG 480
QY 481 AATATTCAAGATGAGTGTGTTTTCATCCATTCATATTCAGCGTACATGACATGAA 540
Db 481 AATGTTCAAGATGCCGTCCTTTCACCCCACTGCAATTCAGCCTACTATGGGACAGAG 540
QY 541 CAGTAACTGCCCTTCTTTGAAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGA 600
Db 541 CAGTAACTGAGTGTCTTTTGAAGTTGGTGTGATGTCATTAAGCGGTGAAAGTTGG 600
QY 601 GATAGACCTCCACCTAGCATCTGCAAGAGGATTCCTGAAATATTCGAAACTCTTGATG 660
Db 601 GACAGGCTCTGCACTTGGGCTCTGCAAGGCTTCTTCAACATTTGGAATCTCTTGTA 660
QY 661 GAAGAAGGCAAGCAGATGTGAATGCTCAAGATAATGAAGACCAATGTCCTCCAT 720
Db 661 GAAGAAGGCAAGCAGATGTGAACGCTCAGGACAAATGAAGACCACTGCTCTGCA 720
QY 721 TTCTGTTCTGATTTGACACCATGATATGATTTAGTATCTGCTGCAAGTGAATGGAA 780
Db 721 TTCTGTTCTGATTTGACACCAATATATAGTAGTACCTGCTCAGAGTGAATAGAG 780
QY 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACACCTGCATCTACAA 840
Db 781 GTCCAGGCTCAGCTATTACATCTATGGTGAACCTCTTGCACCTGCGATGCTACAA 840
QY 841 GGCAAAATTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGATAG 900
Db 841 GGAAATTTGAAGTTGCCAAGGAAATGTCAGGTAACAGGAACTGAAAGTCTGATAG 900
QY 901 GAAACATCTCAGTGAACAGCTTTTCATAGTCTGTGATGCTATGCGAAGAGCATGAC 960
Db 901 GAAACATCTTCAGGAGACAGCTTTTCACAGTCTGTGATGCTATGCGAAGAACATGAC 960
QY 961 CTAGTCAAAATTTCTTGTGATCAGAATGTCTAAAATCAACCAACCAAGGAGGATGG 1020
Db 961 CTGGTCAAAATTTCTTGTGATCAGAATGCTGTGAACATTAACCCAGGAGAGATGG 1020
QY 1021 CACACTGATTAACCTGCTGTGCTACCAAGTCAATTCGCGCTGTTGATGCTTCTACTG 1080
Db 1021 CACACAGGATTTGCACTGCTTGTGCTACCGGCTATTCGCGCTGTTGATGCTTCTACT 1080
QY 1081 GATAATGAGCTGATATGAATCTAGTGGCTTTGATCCAGCAGGCTCTAGTGGTGAATA 1140
Db 1081 GATTAATGCTGAGATATGAATCTTGTGCTTTGATCCAGCAGGCTCTAGTGGTGAATA 1140
QY 1141 GATGAGCAGACATGTTTGTGATGTTGGCTTTGAAAGGAGGATGATGCCATTTGTCACTC 1200
Db 1141 GATGAGCAGACATGTTTGTGATGTTGGCTTTGAGAAAGGACATGATGCCATTTGTCACTC 1200
QY 1201 CTGAAGCATATAGACACCAAGATGAATTCGCTGTAAATGAATTTCTCAGCTGGA 1260
Db 1201 CTGAAGCATATAGACACCCAGGAGGAGCTGCAATGTAACGAATTTCCAGAGCTGGA 1260
QY 1261 GGAGATGGCTCTATGTTGTTTCCATCCCTTGGGGAAGATTAAGAGCATGACAAA 1320
Db 1261 GGAGATGGCTCTATGTTGTTTCCCTTCCCTTGGGCAAGATTAAGAGCATGACAAA 1320
QY 1321 GAGAGGAGATATTTCTTCCCTTAAGAGCTGGATGGCTTCAATTTCCATCTTCTAGCTC 1380
Db 1321 GAGAGGAGATGTTCTTCTTCCCTGAGGCTGAACCTACCCCTCCGCTTCCATCTTCCAACTC 1380

QY 1381 TCAGAAATTCAGTTCCATGAGATTATTGGCTCAGGTTCTTTTGGGAAGTATATAAGA 1440
Db 1381 TCGAAATTCAGTTCCAGAGATTATTCGGCTCGGTTCTTTTGGGAAGTCTATAAAGG 1440
QY 1441 CGATGCAGAAATAAATAGTGGCTATAAAGCTATTCGAGCAATACCTACTGCTCCAAG 1500
Db 1441 CGATGCAGAAATAAATAGTGGCTATAAAGCTATTCGAGCAATACCTACTGCTCCAAG 1500
QY 1501 TCAGATGTGATATGTTTTCGAGAGGTTCCATTTCTGCGAGCTCAATCATCCCTGC 1560
Db 1501 TCAGATGTGATATGTTTTCGAGAGGTTCCATTTCTGCGAGCTCAATCATCCCTGC 1560
QY 1561 GTAAATTCAGTTTGTGGGTGCTTGTGAAATGATCCAGCAGTGTTCCTATGCTCAA 1620
Db 1561 GTGTTTCAGTTTGTGGGTGCTTGTGAAATGATCCAGCAGTGTTCCTATGCTCAA 1620
QY 1621 TACATATCAGGCGGTTCTCTGTTCTCCCTCTCATGACAGAGAGGATTTCTGATTTG 1680
Db 1621 TACATATCAGGCGGTTCTCTGTTCTCCCTCTCATGAAAGAGAGATTTCTGATTTG 1680
QY 1681 CAGTCTAAATTAATTAATTCAGTAGATTTGCAAGGATGAGTACCTTCACAACTG 1740
Db 1681 CAGTCTAAATTAATTAATTCAGTAGATTTGCAAGGATGAGTACCTTCACAACTG 1740
QY 1741 ACACAGCAATATATACATGCTGACTTGAACAGTCAATATTTCTCTATGAGATGG 1800
Db 1741 ACCAGCAATATATACATGCTGACTTGAACAGTCAATATTTCTCTATGAGATGG 1800
QY 1801 CATGCTGTGCTGAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGATGAAGACAAC 1860
Db 1801 CATGCTGTGCTGAGATTTTGGAGAAATCAAGATTTCTGAGTCCCTGGATGAAGACAAC 1860
QY 1861 ATGACAAAACCACTGGGAACCTCGTGGATGGTCTCTGAGGTGTTACGAGTGCACT 1920
Db 1861 ATGACAAAACCACTGGGAACCTCGTGGATGGTCTCTGAGGTGTTACGAGTGCACT 1920
QY 1921 CGGTACACCATCAAGCAGATGTTCTCAGCTATGCTGCTGTCTGCGGAATTTCTCACT 1980
Db 1921 AGATACACCATCAAGCAGTGTGTTCTCAGTTACTCCCTGTCTGTGGAGTCTCTCACT 1980
QY 1981 GGCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGCCAGCAGACATGCTTACCAC 2040
Db 1981 GGAGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGCCAGCAGATATGGGTATCAC 2040
QY 2041 CACATCAGACTCCCATTTGGCTATTCCATTCGCAAGCCCATCTCATCCCTGTGTACGG 2100
Db 2041 CACATCAGACTCCCATTTGGCTATTCCATTCGCAAGCCCATCTCATCCCTGTGTACGG 2100
QY 2101 GGGTGAACCCATGCTCTGAAGGAAGACCCGAATTTCTGAAAGTTGTCATGAAGTTAGAA 2160
Db 2101 GGGTGAACCCATGCTCTGAAGGAAGACCCGAATTTCTGAAAGTTGTCATGAAGTTAGAA 2160
QY 2161 GAGTGTCTGCAACATTCAGTGTCTCTGCTATCAAGTAACAGAGTGGGTCTCTC 2220
Db 2161 GAGTGTCTGCAACATTCAGTGTCTCTGCTATCAAGTAAAGTAAAGTGGGTCTCTC 2220
QY 2221 TCACCTTCTTCTTCTGATTCGCTGTTGACCGGGAGGACCTGGCCGGAGTCAATGG 2280
Db 2221 TCACCTTCTTCTTCTGATTCGCTGTTGACCGGGAGGACCTGGCCGGAGTCAATGG 2280
QY 2281 GCAGCATTAAGAAAGTGGTTTGAATTTGAGTTGGAGTATGCTTAAATGCAAGGTCCTATGCTGT 2340
Db 2281 GCAGCATTAAGAAAGTGGTTTGAATTTGAGTTGGAGTATGCTTAAATGCAAGGTCCTATGCTGT 2340
QY 2341 TTGTCCCAAGAGTGTGGAACAAATTTCTCTCAAGGCTCTCTTTGGAGGAGATGAAGA 2400
Db 2341 TTGTCCCAAGAGTGTGGAACAAATTTCTCTCAAGGCTCTCTTTGGAGGAGATGAAGA 2400
QY 2401 AGTCTTCAATACACACCATTTGACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460
Db 2401 AGTCTTCAATACACACCATTTGACAAATATGGCTATGATCCGATCCCATGAGCTCAATG 2460
QY 2461 CATTTTCATTTCTTGGCCAAATATGATGACGCTTTGAGGACAGCAGC 2505

Db 1681 TACATTTTCAGAGGCTCCCTGTTCTCCCTGCTTCATGAACAGAGAGAAATCTTGACTTG 1740
Qy 1681 CAGCTAAATTAATTTTCAGTAGATGTTGCCAAGGCGATGAGTACCTTCACAACTG 1740
Db 1741 CAGTCTAAATTAATTTTCAGTAGATGTTGCCAAGGCGATGAGTACCTTCACAACTG 1800
Qy 1741 ACACAGCAATTAATTCAGTAGATGTTGCCAAGGCGATGAGTACCTTCACAACTG 1800
Db 1801 ACCAGCAATTAATTCAGTAGATGTTGCCAAGGCGATGAGTACCTTCACAACTG 1860
Qy 1801 CAGTCTGAGTGGAGATTTTGGAGATCAAGATTTTTCAGTCTCTGATGAGTGGAG 1860
Db 1861 CAGTCTGAGTGGAGATTTTGGAGATCAAGATTTTTCAGTCTCTGATGAGTGGAG 1920
Qy 1861 ATGACAAACAACTGGAGACCTCCGTTGGATGGCTCCCTGAGTGTTCACGAGTGCAC 1920
Db 1921 ATGACAAACAACTGGAGACCTCCGTTGGATGGCTCCCTGAGTGTTCACGAGTGCAC 1980
Qy 1921 CGGTACACCAATCAAGAGATGTTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 1981 AGATACACCAATCAAGAGATGTTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Qy 1981 GGGGAAATTCATTCGCTCATCTCAAGCAGAGGCTGGCGAGCAGACATGGCTTACAC 2040
Db 2041 GGGGAAATTCATTCGCTCATCTCAAGCAGAGGCTGGCGAGCAGACATGGCTTACAC 2100
Qy 2041 CACATCAGACCTCCCATTTGGCTATTCATTCACAGCCCATATCATCTCTGCTGATAC 2100
Db 2101 CACATCAGACCTCCCATTTGGCTATTCATTCACAGCCCATATCATCTCTGCTGATAC 2160
Qy 2101 GGGTGGAGCAGTGTCTGAGAGAGAGCCGATTTTCTGAGTGTGCTGCTGCTGCTGCTGCT 2160
Db 2161 GGGTGGAGCAGTGTCTGAGAGAGAGCCGATTTTCTGAGTGTGCTGCTGCTGCTGCTGCT 2220
Qy 2161 GAGTGTCTGCTGCAACATGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 2221 GAGTGTCTGCTGCAACATGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Qy 2221 TCACCTTCT 2280
Db 2281 TCACCTTCT 2340
Qy 2281 GCACATTAAGAGAGTGGTTCGAATTTGGAATATGCTTAAATGCAAGTGTCTATGCTGCT 2340
Db 2341 GCACATTAAGAGAGTGGTTCGAATTTGGAATATGCTTAAATGCAAGTGTCTATGCTGCT 2400
Qy 2341 TTGTCCAAAGTGTGACCAATATTCCTCTCAAGGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
Db 2401 TGGTCCAAAGTGTGACCAATATTCCTCTCAAGGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2460
Qy 2401 AGTCTTCAATACACACCAATGACCAATATGCTTAAATGCAAGTGTCTATGCTGCTGCT 2460
Db 2461 AGCACCAGTATTAAGTGTGACCAATATGCTTAAATGCAAGTGTCTATGCTGCTGCTGCT 2520
Qy 2461 CATTTTTCATTTCTTGGCAATATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2505
Db 2521 CACCTTCACTCCCGCAAGAGCAGACAGCACTTTTGGAGCAGCAAC 2565

RESULT 8

US-10-626-173-7
; Sequence 7, Application US/10626173
; Publication No. US20040110232A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL DARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/10/626,173
; PRIORITY FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: US/09/947,199A
; PRIOR APPLICATION NUMBER: US 60/111,938
; PRIOR FILING DATE: 1998-12-11

; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3026
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(2565)
US-10-626-173-7

Query Match 75.5%; Score 1892.2; DB 17; Length 3026;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 1 ATGGGAAATTAATAATCTATAGACCAACCCAACTTGTCTGATGAATGGAAGAAAAGTTC 60
Db 61 ATGGGAAATTAATAATCTATAGACCAACCCAACTTGTCTGATGAATGGAAGAAAAGTTC 120
Qy 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGTGAAGTCAAGTCAAGAAAAA 120
Db 121 AGTGAATCTTACGCTATTAATCAAGAAAGCTGGAGATACTGCAGATCAAGAAAAA 180
Qy 121 GAAGTCAAGAACTAAGGAATATATTTGGCTCTGATGAAGCTTCAGTAAAGTCAATTTA 180
Db 181 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCTTCAGTGAAGTCAATTTA 240
Qy 181 AATTACGGCACTGAAATGGGCTGCTCTACTCTCAATTTATGTTGATTTGAGGCGAAG 240
Db 241 AATTACGGCACTGAAATGGGCTGCTCTACTCTCAATTTATGTTGATTTGAGGCGAAG 300
Qy 241 AATTCACATATTCGAATCTTATGTTGAAAGGGCTCCGCCATCTCGACTGCAAGAAAT 300
Db 301 AATTCACATATTCGAATCTTATGTTGAAAGGGCTCCGCCATCTCGACTGCAAGAAAT 360
Qy 301 GGAATTTACAGCTTGAATTTAGCAGTTTACAGAGTAATGAGAAATGATGATCTCTCTG 360
Db 361 GGGTTTTCCAGCTCTGACCTGGCGCTTTTACAGAGCAGCCCGAACTTATCACTTCAC 420
Qy 361 CTTCACAGTGGAGCTGATATACAGAGGTTGATACGGTGGCTCACTGCGCTCCCATATT 420
Db 421 TTGCACAGCGAGCAGATGTTTACAGAGTGGATACGGTGGCTCACAGCCCTCCACATA 480
Qy 421 GCTACAAATAGCTGGCCACCTTAGAGGCTGCTGATGCTGTTTGCACATGAGAGCTAATGTC 480
Db 481 GCTGCAATAGCTGGACACACCCAGAGGCTGCAGAAAGTCTGCTACAAATGCGGCGCAATG 540
Qy 481 AATATTCAGATGCACTTTTTCCTCCATTCATATTCGAGCTGCTGCTGCTGCTGCTGCT 540
Db 541 AATGTTCAAGATGCGCTTCTTTCACCCACATTCGACATTCGAGCTTATGCGGCAAG 600
Qy 541 CAGGTAACCTGCGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTGA 600
Db 601 CAGGTAACCTGCGCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTGG 660
Qy 601 GATAGACCCCTCCACCTAGATCTGCAAAAGAGATTTTGAATTTGCAAACTCTTGATG 660
Db 661 GACAGGCTCTGACCTGGCTCTGCAAAAGAGATTTTGAATTTGCAAACTCTTGATG 720
Qy 661 GAAGAGGCGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACATGCTCCCATCCAT 720
Db 721 GAAGAGGCGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACATGCTCCCATCCAT 780
Qy 721 TTCTGTTCTCGATTTGGACCAACATATAGTAAAGTATCTCTGCTGCAAGTGAATGGA 780
Db 781 TTCTGTTCTCGATTTGGACCAACATATAGTAAAGTATCTCTGCTGCAAGTGAATGGA 840
Qy 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGCTGCTGCTGCTGCT 840

Db 841 GTCCAGCCTCAGCTCAATTAACATCTATGGTGACACTCCTTTGGACCTGGGCATGCTACAAT 900
QY 841 GGCAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGAAACAGAAAGTCTGACTAAG 900
Db 901 GGAATTTTGAAGTTGCCAAGGAATTTGCCAGGTAAACAGAACTGAAGTCTGACTAAG 960
QY 901 GAAACATCTTCAGTGAACAGCTTTTCATAGTCTGTTGACCTATGCTATGCTGCAAGACATTTGAC 960
Db 961 GAAACATCTTCAGTGAACAGCTTTTCATAGTCTGTTGACCTATGCTATGCTGCAAGACATTTGAC 1020
QY 961 CTAGTCAAAATTTCTTTGATCAGAAATGCTATAAATCAACATCAACCAAGGAAGGATGGG 1020
Db 1021 CTGGTCAAAATTTCTTTGATCAGAAATGCTATAAATCAACATTAACCAAGGAAGATGGG 1080
QY 1021 CACATGGATTAACATCTGCTGTGCTACCAAGGTCAATTCGCTGTTGCTTCTTACTG 1080
Db 1081 CACACAGGATTCATCTGCTGTGCTACCAAGGTCAATTCGCTGTTGCTTCTTACTG 1140
QY 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTTGATCCCAAGGTCTAGTGGTGAAGAAA 1140
Db 1141 GATAATGGTGCAGATATGAATCTTGTGCTGCTGATCCCAAGGTCTAGTGGTGAAGAAA 1200
QY 1141 GATGAGCAGACATGTTGATGGCTTTATGAAAAGGCGATGATGCCATTTGTCACATC 1200
Db 1201 GATGAGCAGACATGTTGATGGCTTTATGAAAAGGCGATGATGCCATTTGTCACATC 1260
QY 1201 CTGAAGCATTTAAGAGACCAACAAGATGAATTCGCTGTTGATGAATTTCTCAGCCTGGA 1260
Db 1261 CTGAAGCATTTAAGAGACCAACAAGATGAATTCGCTGTTGATGAATTTCTCAGCCTGGA 1320
QY 1261 GGAGATGGCTCTATGCTCTGTTCCATCACCTTGGGGAAGATTAAGACATGACAAAA 1320
Db 1321 GGAGATGGCTCTATGCTCTGTTCCATCACCTTGGGGAAGATTAAGACATGACAAAA 1380
QY 1321 GAGAGGCGAGATTTCTCCTCTGAGGGCTGAACTACCTCCGCTTCCATCTCCAACTC 1380
Db 1381 GAGAGGCGAGATTTCTCCTCTGAGGGCTGAACTACCTCCGCTTCCATCTCCAACTC 1440
QY 1381 TCAGAAATTTGAGTTCCATGAGATTAATGGCTCAGGTTCTTTTGGGAAGATATAAGGA 1440
Db 1441 TCCGAATTCGAGTTCCACGAGATTAATGGCTCGGTTCCCTTTGGGAAGATCTATAAGGG 1500
QY 1441 CGATCCGAATTAATTAAGTGGCTATAAAGCTTATCAGGCAATACCTACTCTCTCAAG 1500
Db 1501 CGATCCGAATTAATTAAGTGGCTATAAAGCTTATCAGGCAATACCTACTCTCTCAAG 1560
QY 1501 TCAGATGTGGATATGTTTTCGAGAGGTGTCCATCTCTGCCAGCTCAACACCTCTGC 1560
Db 1561 TCAGAGCTGGATATGTTTTCGAGAGGTGTCCATCTCTGCCAGCTCAACACCTCTGC 1620
QY 1561 GTAATTCAGTTTGGGTGCTGTTGTAATGATCCAGCAGGTTGCTCATTTGCTCACTCAA 1620
Db 1621 GTGGTTCAGTTTGGGTGCTGCTGATGACCCAGTCAGTTGCTCATTTGCTCACTCAG 1680
QY 1621 TACATATCAGGGGTTCTGTTCTCCTCCTTCATGAGAGAGGATTTCTTGAATTTG 1680
Db 1681 TACATTTGAGAGGCTCCTGTTCTCCTGCTTCATGAGAGAGATTTCTTGAATTTG 1740
QY 1681 CAGTCTAAATTAATTTGAGTATGTTTGCAGAGGATGAGTACCTTCAACACCTG 1740
Db 1741 CAGTCTAAATTAATTTGAGTATGTTTGCAGAGGATGAGTACCTTCAACACCTG 1800
QY 1741 ACACAGCAATTTATACATGCTGCTTGAAGTCAACATATTTCTCTATGAGGATGG 1800
Db 1801 ACCAGCAATTCATACCCGAGCTGAACAGCCACCAATTTCTGCTATGAGGATGGC 1860
QY 1801 CATGCTGTGGGCGAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
Db 1861 CATGCTGTGGGCGAGATTTTGGAGAAATCAAGATTTCTGAGTCTCTGGATGAAGACAAC 1920
QY 1861 ATGACAAAACCTGGGAACTCGTTGGATGGCTCTGAGGTGTTTCAAGAGTCACT 1920
Db 1921 ATGACAAAACCTGGGAACTCGTTGGATGGCTCTGAGGTGTTTCAAGAGTCACT 1980

QY 1921 CGGTACACCATCAAGACAGATGCTTTTCAGCTATGCTCTGCTGTGGGAAATTTCTCACT 1980
Db 1981 AGATACACCATCAAGGCTGATGCTTTTCAGTTACTCTCCTGTGTGTGGGAGTCTCTCACT 2040
QY 1981 GGGCAAAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGGACAGACATGCTGCTTACCAC 2040
Db 2041 GGAGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGCGGACAGATATGGGTATCAC 2100
QY 2041 CACATCAGACTCCCATTTGGCTATTCCATTCCCAAGCCCATATCATCTCTGCTGATACA 2100
Db 2101 CACATCAGAGCGGCCATTCGCTATTCCATCCCAAGCCCATCTCATCCCTGCTGATACGG 2160
QY 2101 GGGTGGAAACCATGCTCTGAAGCAAGACCGCAATTTTCTGAAGTTCTCATGAAGTTAGAA 2160
Db 2161 GGCTGGAAATGCTATGCTCTGAAGCAAGACAGAGTTCTCTGAAGTCTGTAGCAAACTGGAG 2220
QY 2161 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGCAATCAAGTAACAGCAGTGGGTCTCTC 2220
Db 2221 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGCAATCAAGTAACAGCAGTGGGTCTCTC 2280
QY 2221 TCACCTTCTCTCTCTCTCTGATTCGCTGTAACCGGGAGGACCTGGCCGAGTCAATG 2280
Db 2281 TCACCTTCTCTCTCTCTCTGATTCGCTGTAACCGGGAGGACCTGGCCGAGTCAATG 2340
QY 2281 GCAGCATTAAGAAAGTCTCTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCT 2340
Db 2341 GCAGCATTAAGAAAGTCTCTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCT 2400
QY 2341 TTGTTCCAAAGTCTGCAACATTAATTCCTCTCAAGGTCTCTCTTGGAGAGATGAAGA 2400
Db 2401 TGGTCCCAAGTGTGGAACACACTCTTAATCCGGGCTCTCTTGGAGAGATGAATAGG 2460
QY 2401 AGTCTTCAATACACACCCATTGCAAAATATGGTATGATCGATCCCATGAGTCAATG 2460
Db 2461 AGCACCAGATTTCAACTGTTGACAAATACGCTATGCTCTGATCCCATGAGCTGAG 2520
QY 2461 CATTTTCATTTCTGCGCAATATAGTACAGCTTTTGGAGCAGCAGC 2505
Db 2521 CACCTTCACTCCCGCAAGACGACGACCACTTTTGGAGCAGCAAC 2565

RESULT 9

US-09-833-381-1149
; Sequence 1149, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1: Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1149
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1149

Query Match 17.1%; Score 427.4; DB 9; Length 616;
Best Local Similarity 81.8%; Pred. No. 4.3e-119;
Matches 505; Conservative 0; Mismatches 111; Indels 1; Gaps 1;
QY 291 GACAAGAAATTCGATTTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCAAGATTCAT 350
Db 1 GACGAGAAATGGGTTTTCCAGCTCTGACCTGCGCGCTTTACAAGGACAGCCCGAACTTAT 60
QY 351 CACTTCTCTCTTCAAGTGGAGCTCATATACAGAGGTTGATACGGTGGCTCTCACTGC 410
Db 61 CACTTCACTGTTCAGAGCGGAGCAGATGTTTCAAGCAAGTGGATACGGTGGCTCTCACAGC 120

Qy 411 CCTCCATATTGCTACATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTTTGCAACATGG 470
Db 121 CCTCC-CATAGCTGCAATAGCTGGACACCCAGAGGCTGCAGAAGTGTCTGCTGCAACATGG 179
Qy 471 AGCTAATGTCATATTCAAGATCGAGTTTTTTTTCACCTCCATTCATATTCGACGCTACTA 530
Db 180 GGCCAAAGCTGAATGTTCAAGATGCCGTCTTCTTCACCCACCTGCAATTCGACGCTACTA 239
Qy 531 TGGACATGAACAGAGTAACTCGCCCTCTCTTTTGAATTTGGTGTGATGTAAATGTAAGTGG 590
Db 240 TGGCACGACGAGTAACCACTGTCCTTTTGAAGTTTGGTGTGATGTCAATGTAAGCGG 299
Qy 591 TGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGATTTCTTGAATTTGCAAA 650
Db 300 TGAAGTTGGGACAGGCTCTGCACCTGGCCCTCTGCAAAAGGCTTCTTCAACATTTGTAA 359
Qy 651 ACTCTCATGGAAGAGGACGAGCAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGT 710
Db 360 ACTCTGTAGAGAGGAGCAGAGCATGTGAACGATGTGAACGCTCAGGACNATGAAGACCACT 419
Qy 711 CCACATCCATTTCTGTCTCGATTTGGACACCATGATATAGTTAAGTATCTCTGCTGCAAG 770
Db 420 CCCTCTGCACTTCTGTCTCGATTTGGACACCAATATAGTGAGCTACCTGCTCCAGAG 479
Qy 771 TGAATTTGGAAGTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACTGGC 830
Db 480 TGACTTAGAGTCCAGGCTCACGTCTAATACATCTATGTTGACATCTCTTTCGACCTGGC 539
Qy 831 ATGTACAATGGCAAAATTTGAAGTTGCCAAGGAAATCATCCAATATCAGGAACAGAAAG 890
Db 540 ATGTACAATGGAAATTTTGAAGTTGCCAAGGAAATTTCCAGGTAAACAGGAACGTAAAG 599
Qy 891 TCTGACTAAGGAAACA 907
Db 600 TCTGACTAAGGAAACA 616

RESULT 10

US-09-833-381-1150
; Sequence 1150, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1150
; LENGTH: 736
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(736)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1150

Query Match 16.6%; Score 414.6; DB 9; Length 736;
Best Local Similarity 81.2%; Pred. No. 4e-115;
Matches 480; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
Qy 265 TTGAAGGGCTCCGCCCATCTCGACTGACAGAAATGGATTTACAGCTTGATTTAGCA 324
Db 1 TTAAAGGGCTCCGTCATCCAGACTCAGAGAAATGGGTTTCCAGCTCTGACCTGGCC 60
Qy 325 GTTTACAGGATAATGAGAAATTCATCTCTCTGCTTCACAGTGGAGCTGATATACAG 384
Db 61 GTTTACAGGACACCCCGAACATATCACTTCACTGTTCACAGCGGAGCAGATGTTTCA 120

Qy 385 CAGTTTGGATACGGTGGCCTCACTGCCCTCCATATTCCTACAAATAGCTGGCCACCTAGAG 444
Db 121 CAAAGTGGGATACGGTGGCCTCCAGCCCTCCACATAGCTGCAATAGCTGGACACCCAGAG 180
Qy 445 GCTGCTGATGCTGTTTGCACATGGAGCTAATGTCAATATTCAAGATGCGAGTTTTTTTC 504
Db 181 GCTGCGAAGTGTGCTTCAACATGGGCGCAATGTGAATGTTTCAAGATGCGGCTCTTCTTC 240
Qy 505 ACTCCATTTGCATATTCAGCGCTACTATGGACATGAACAGGTAACTCGGCTTTCTTTGAAA 564
Db 241 ACCCCACTGACATTGCAGCCTACTATGGCCACGAGCAGGTAAACAGTGTCTCTTTTGAAG 300
Qy 565 TTTGGTGTCTGATGAATGAATGAGTGGTGAAGTGGAGATAGACCCCTCCACCTAGCATCT 624
Db 301 TTTGGTGTCTGATGTCAATGTAAAGCGGTGAAGTGGGACAGGCGCTCTGACCTGGCCCTCT 360
Qy 625 GCAAAAGGATTTCTTGAATATTTGCAAACTCTTTGATGGAAGAGGAGGAGCAAGCAGATGTG 684
Db 361 GCAAGGGCTTCTTCAACATTGTGAAACTCTCTGGTAGAAGAGGAGCAAGCAGATGTG 420
Qy 685 AATGCTCAAGATTAATGAAGACCAATGTCCACCTCCATTTCTGTTCTCGATTTGGACACCAT 744
Db 421 AAGCGCTCAGACCAATGAAGACCAACGCTCCCTCTGCACTTCTGTTCTCGATTTGGACACCAT 480
Qy 745 GATATAGTTAAGTATCTGCTGCAAACTGATTTGGAAGTTCAACCTCATGTTCTTAAATATC 804
Db 481 AATATAGTAGCTACTGCTGCTCCAGAGTGACTTANAGGTCACGCTCAGCTCATTAACATC 540
Qy 805 TATGGAGATACCCCTTTACACCTGGCATGCTCAATGGCAAAATTTGAAGTT 855
Db 541 TATGGTGACACTCCTTTGCACTGCGCATGCTCAATGGAAATTTTGAANTT 591

RESULT 11

US-09-833-381-1151
; Sequence 1151, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1151
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(740)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1151

Query Match 16.5%; Score 413.2; DB 9; Length 740;
Best Local Similarity 79.0%; Pred. No. 1.1e-114;
Matches 498; Conservative 0; Mismatches 131; Indels 1; Gaps 1;
Qy 1 ATGGGAAATTAATAATCTAGACCAACCCAAACTTGTACTGATGAATGGAGAAAAAGTC 60
Db 73 ATGGGAAATTAATAATCTAGACCAACCCAACTGTTCTGATGAATGGAGAAAGAAAGTT 132
Qy 61 AGTGAATCATATGTTATCAATATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 120
Db 133 AGTGAATCTTACGCTATTATCATAGAAAGGCTGGAGGATAA CTTGCAGATCAAGAAAT 192
Qy 121 GAACGTGACAGAACTAAGGAATATATTTGGCTCTGTATGGAAGCCTTCAGTAAAGTCAATTTA 180
Db 193 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGTATGAAGACCTTCAGTGAAGTCAGTTTA 252

```
QY 181 AATTACCGCACTGAAAAAGGGCTGTCTCTACTTATTTATGTTGCAATTTGTGAGGCAAG 240
Db 253 AATTACCGCACAGACGGTGGCTGTCCCTGTCTACACCTCTGCTGTGTCTGTGGGCAAC 312
QY 241 AATACACATATTCGAACCTTTATGTTGAAAGGGCTCGCCCATCTCGACTGACAGAAAT 300
Db 313 AAGTCACATATCCGTGGCCCTTATGTTTAAAGGGTCCGTCCATCCAGACTGACGAGAAAT 372
QY 301 GGATTTACAGCCCTTCATTTAGCAGTTTACAGAGATAATGAGAAATGATCACTTCTCTG 360
Db 373 GGGTTTCCAGCTCTGCACCTGGCCGTTTACAGAGACGCCGGAACCTTATCACTTCACTG 432
QY 361 CTTTAC-AGTGAGCTGATATACAGCAGGTTTGGATACGGTGGCCCTCACTGCCCTCCATAT 419
Db 433 TTGCACAAGCGGAGCAGATGTTACAGAAAGTGGATACGGTGGCCCTCACAGCCCTCCACAT 492
QY 420 TGCTACATAGCTGCCACCTAGAGGCTGCTGATGCTGTGCAACATGAGCTAATGT 479
Db 493 AGCTGCAATAGCTGGACACCCAGAGGCTGCANAAAGTGTGNTGCAACATGGGCCAACGT 552
QY 480 CAATATTCAAGATGCAAGTCTTTTTCATCTCCATTTGCATATTTGCAGCGTACTATGGACATGA 539
Db 553 GAATGTTCAAGATGCGCTTCTTTCACCCCACTGCACATTTGCAGCCTACTATGGCACGA 612
QY 540 ACAGTAACTGCCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGTGAAAGTTGG 599
Db 613 GCAGGTAACCAAGTGTCTCTTTTGAAGTTTGGNGGTGATGTCNATGTAAGCGGTNAAGTTGG 672
QY 600 AGATAGACCCCTCCACCTAGCATCTGCAAA 629
Db 673 GNAACGGCTCGGCCCTTGGCCNCGNAAA 702
```

RESULT 12

```
US-09-833-381-1147/c
; Sequence 1147, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1147
; LENGTH: 304
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(304)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1147
```

```
Query Match 7.4%; Score 184.2; DB 9; Length 304;
Best Local Similarity 79.6%; Pred. No. 5.1e-45;
Matches 242; Conservative 0; Mismatches 59; Indels 3; Gaps 2;

QY 266 TGAAGGGCTCGGCCATCTCGACTGACAGAGAAATGGAATTACAGCC--TTGCAATTTAGC 323
Db 304 TAAAGGGCTCGTCCATCCAGACTGACGAGAAATGGGTTTCCAAAGCTTTGCAACCTGCC 345
QY 324 AGTTTACAGGTAATGAGAAATGATCAG--TTCTCTGCTTCAAGTGGAGCTGATATAC 382
Db 244 GGTTTACAGAGACAGCCCGGAATTTATCACTTTCACTGTTGACAGCGGAGCAGATGTTTC 185
QY 383 AGCAGGTTGGATACGGTGGCCCTCACTGCCCTCCATATTTGCTACAATAGCTGGCCACCTAG 442
Db 184 AGCAAGTGGGATACGGTGGCTCCACAGCCCTCCACATAGCTGCAATAGCTGGACACCCAG 125
```

```
QY 443 AGGCTGCTGATGTCTGTTTGCACATGGAGCTAATGTCATATTCAGATGCAAGTCTTTT 502
Db 124 AGGCTGCAGAAAGTCTGATGCAACATGGGGCCAAACGTAATGTTCAAGATGCGTCTTCT 65
QY 503 TCACTCCATTCATATTTGAGCGTACTATGGACATGAAAGGTAACCTCCGCTTCTTTGA 562
Db 64 TCACCCCACTGCACATTCGAGCNTACTTGGGCACGAGCGGTAACCACTGTCCTTTGA 5
QY 563 AATT 566
Db 4 AGTT 1

RESULT 13
US-10-424-599-53404
; Sequence 53404, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 53404
; LENGTH: 1953
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_19237C.1
US-10-424-599-53404
```

```
Query Match 4.4%; Score 109.6; DB 13; Length 1953;
Best Local Similarity 51.3%; Pred. No. 1.3e-21;
Matches 390; Conservative 0; Mismatches 349; Indels 21; Gaps 5;

QY 1405 ATTGGCTCAGGTTCTTTTGGAAAGTATATATAAGGACGATGACAGAAATAAATACTGGCT 1464
Db 884 ATTGCATCTGGGTCAATATGTTGAACATTTTAAAGGTGTATATTGTAGCCAGGAAGTGGCC 943
QY 1465 ATAAACGTTATCGAGCCATACCTACTGCTCCCAAGTCAGATGTGATATGTTTGGCGA 1524
Db 944 ATCAAGTTCTCAAGGCTGACCATGTAACCTGAAATTCAGAGAGA---GTTCGACAG 1000
QY 1525 GAGGTGTCATTTCTCTGCCAGCTCAATCATCCCTGCGTAATTCAGTTGTGGTGGTCTTGC 1584
Db 1001 GAAGTCTATATCATGAAAGGTTTCGACACAAAGATGTTGTACAAATTCATAGGAGCTTGT 1060
QY 1585 TTGAATGATCCAGCCAGTTTGCATGTCCTCAATACATATATCAGGGGTTCTCTGTC 1644
Db 1061 ACCAAGTCCCAAGCT---TTATGCAATAGTAAAGAAATTTATGTCGTGGTGGAAAGTGTGAT 1117
QY 1645 TCCCTCTCTCATGACGAGAGAGGATCTTTGATTTTGCAGTCTAAATTAATTTATTCAGATGA 1704
Db 1118 GACTACCTACATAAGCAGAGAGGCTTTTAAATTTCTTCACTGCTCAAGTAGCAAT 1177
QY 1705 GATGTTGCCAAAGCATGAGTACCTTTCACACCTGACAGCCCAATATACATGTCAG 1764
Db 1178 GATGTTTCCAAAGGAATGAATCTTTCACCA-----ACATAATATATATCCATAGAGAC 1231
QY 1765 TTGAACAGTCACAAATATCTTCTATGAGGATGGGCACTGCTGTGTGGTGGCAATTTTGA 1824
Db 1232 TTGAAGCTCCCAACCTTTTGTGATGAAATTTGACTGTAAGGTGCTGTGATTTGGG 1291
QY 1825 GAATCAAGATTTTACAGTCTCTGGATGAAGACCAATGACAAAACAACCTGGGAACCTC 1884
Db 1292 GTTCTAGAGTTAAA-----GCTCAATCTGGCGTCACTGCTGCAAAACTGGAAACATAT 1345
```

QY 1885 GGTGGATGGCTCTGAGGTTTCACGAGTGCACCTCGGTACACCATCAAGCAGATGTC 1944
Db 1346 CGATGGATGGCTCTGAGGTTTAGAACA---CAAGCGTAGATCACAAGGCTGATGA 1402
QY 1945 TTCAGTATGCTCTGTGTCTGTGGGAAATTCACCTGGCGAAATTCATTCGCTCATCTC 2004
Db 1403 TTTAGTTTGGAAATGTTTATGGAGTTGCTCACCGGAAAGCTTCATACGAATATTA 1462
QY 2005 AAGCCAGCGGCTCGGCAGCAGACATGGCTTACCACCATCAGACCTCCATGGCTAT 2064
Db 1463 ACCCCCTACAGGAGCTATAGGAGTGGTTTCAAAGGGTTGCGACCCACCATCCGAG 1522
QY 2065 TCATTCGCCAAGCCATCATCTCTGCTGATACGAGGGTGAACGCATGCTCTGAAGA 2124
Db 1523 AACACTCATCAAAGTTTGGAGCTCTTGAGAGGCTCTGGCAGCAAGATCCTACATTG 1582
QY 2125 AGACCCGAATTTCTGAAGTTGTCATGAAGTTAGAAGT 2164
Db 1583 AGACCTGATTTCTCGGAATATCGAGATCCTCGACAGT 1622
RESULT 14
US-10-424-599-129156
; Sequence 129156, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Zhou Yihua
; APPLICANT: Kovalic David K
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 129156
; LENGTH: 3238
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87633C.1
US-10-424-599-129156
Query Match 4.3%; Score 107.4; DB 13; Length 3238;
Best Local Similarity 52.4%; Pred. No. 8.8e-21;
Matches 387; Conservative 0; Mismatches 331; Indels 21; Gaps 6;
QY 1406 TTGGCTCAGGTTCTTTTGGGAAGTATATAAGGACGATCGAGATAAATAGTGGCTA 1465
Db 1615 TTGGCTCTGGTCAATTTGGTGCTTTGACAGGTACATATGTAGTCAAGATGGCTA 1674
QY 1466 TAAACGTTATCGAGGCAATACCTACTGCTCCAAAGTCAGATGGATATGTTTGGCCAG 1525
Db 1675 TCAAA-GTCTTAAGCTGAGCGCAATAGTACAGATATGCTGAGG--GAGTTTGCACAGG 1731
QY 1526 AGGTGTCATCTCTGCCACTCAATCATCCCTGCGTAATTCAGTTTGGGTGCTTGT 1585
Db 1732 AAGTTTATATCATGAGGAAGATTCGACACAAGAAATGTTGTCAGTTCAITGGCGCATG-- 1789
QY 1586 TGAATGATCCAGCGAGTTTGGCAATGTCTCAATACATATCAGGGGTTCTCTGTTCT 1645
Db 1790 -TACTAGGCCCCAAATCTTTGTTATGTTACTGAGTTTATGCTAGGGAAGCTTATG 1848
QY 1646 CCCTCTCTATGACGAAGAGATTTCTGATTTGGAGTCTAAATTAATTTAGCAGTAG 1705
Db 1849 ACTTCTGCAACAACAAGAGGTGTATTTAAGCTTCCATCTTTGCTTAAAGTAGCAATG 1908
QY 1706 ATGTTGCCAAGGCTGAGTACCTTCACAACTGACAGGCAATATACATCTGACT 1765
Db 1909 ATGTTTCAAGGGAATGAATCTTTTGACCA-----AAATATATTAATTCACAGGACC 1962
QY 1766 TGAACAGTCACAATATCTTCTCTATGAGGATGGGATGCTGTGGTGGCAGATTTTGGAG 1825

Db 1963 TCAGACTGCCATCTTCTGATGGATGAATGAAGTGTCAAGGTTGCTGATTTGGGG 2022
QY 1826 AATCAAGATTTTACAGTCTCTGGATGAAGACAAACATGACAAAAAACCTGGGAACCTCC 1885
Db 2023 TTGCCAGGGTGCAA-----ACTCAGTCTGGAGTGATGACAGCTGAAACTGGAACATACC 2076
QY 1886 GTTGGATGGCTCTGAGGTTTCAAGCAGTGCACCTGGGTACACCATCAAGCAGATGTCT 1945
Db 2077 GTTGGATGGCTCTGAGG---TCATTTGAACACAACCATATACCAAGAGGAGATGTT 2133
QY 1946 TCAGCTATGCTCTGTCTGTGGGAAATTCACCTGGCGAAATTCATTCGCTCATCTCA 2005
Db 2134 TCAGTTTGGAAATAGCTCTTTGGAGCTTTTAACTGGAGAACTGCTTACTCTTCGCTGA 2193
QY 2006 AGCAGCGGCTCGGCAGCAGACATGGCTTACCACCATCAGACCTCCCATTTGGCTATT 2065
Db 2194 CCCATTACAGCAGAGTTGGCGTGTGCGAAGGGCTCGGGCTTACAATTCACAAA 2253
QY 2066 CCATTTCCCAAGCCCATATCATCTCTGCTGATACGAGGGTGGAAACGATGTCTCTGAAGGAA 2125
Db 2254 ATAGCACCACCAAGACTTTCTGAACTTTCTCAGCGGTGCTGGCAACAAGATCCAAACGCAA 2313
QY 2126 GACCCGAATTTTCTGAAGT 2144
Db 2314 GACCAAAATTTCTCTGAAAT 2332
RESULT 15
US-10-425-114-22955
; Sequence 22955, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22955
; LENGTH: 2207
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3592-081-B1_FLI
US-10-425-114-22955
Query Match 4.0%; Score 100.2; DB 13; Length 2207;
Best Local Similarity 52.1%; Pred. No. 1e-18;
Matches 328; Conservative 0; Mismatches 283; Indels 18; Gaps 4;
QY 1516 TTTTGGCGAGAGTGTCATCTCTGCCAGCTCAATCATCCCTCGGTAAATTCAGTTTGTG 1575
Db 1133 TTTGCTCAGGAGTATATATATATGAGGAAGGTCGCTCACAAGAATGTTGTGCAATTTAT 1192
QY 1576 GGTGCTTGTGTAATGATCCAGCCAGTTTGGCATTTGTCATTCACATACATATCAGGGGT 1635
Db 1193 GGGCTTGGACTTAACCCCAACCTATGT---ATAGTCACAGAATTTATGTCGGTGA 1249
QY 1636 TCTCTGTCTCCCTCTTCATGAGCAGAAGAGATTTCTTGATTTGCAGTCTAAATTAAT 1695
Db 1250 AGTGTGTATGATTACCTCCATAAACAAGGTGTTTTCAAACTTCTGCTTTAGTTGA 1309
QY 1696 ATTGCAGTAGATCTTCCCAAGGCATGGAGTACCTTCACAACCTGCACAGCCATTATA 1755
Db 1310 GTTGCACAGGATGTGCAAAAGGCATGAGCTACTTACACCAAGATAAT-----ATTATT 1363

QY	1756	CATCGTACTTGACAGTCACAATATTTCTCTATGAGGATGGCATGCTGTGTGGCA	1815
Db	1364	CATCGTATTTGAAACTGCAACCTTCTGATGGATGAATGGACTGTTAAGTTGCT	1423
QY	1816	GATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACACATGACAAACCAACCT	1875
Db	1424	GATTTTGGTGTGGCAGCTGTTAAAGCTCAATCTGGAGTT-----ATGACTGCAGAAACT	1477
QY	1876	GGGAACCTCCGTTGGATGGCTCCTGAGGTGTTACGAGTGCACCTCGGTACACCATCAA	1935
Db	1478	GGTACTTACCGTTGGATGGCCCGCAGAGGTCAATTGAGCA---CAAACCTATGATCAAG	1534
QY	1936	GCAGATGTCTTCAGCTATGCTGTGTCTGTGGGAAATTTCTCACTGGCGAAATTCATTTC	1995
Db	1535	GCTGATGTTTTAGTTTTTGGAAATTTGATGTGGGAACTACTCACAGGGAAGATTCCTTAT	1594
QY	1996	GCTCATCTCAAGCCAGCGGCTGGGCGACGACATGGCTTACCAACCATCAGACCTCCC	2055
Db	1595	GAGTACCTGACTCCACTCAAGCCGCTGTTGGTGTGTCAGAGGGATTACGGCCTACA	1654
QY	2056	ATTGGCTATTCATTCCCAAGCCCATATCATCTCTGTGTATACGAGGTTGGAACGCATGT	2115
Db	1655	ATCCAAAGCATCTACGCAAGCTTTCGAGTTCTTCAGAAATGCTGGCAGCAGGAC	1714
QY	2116	CCTGAAGGAAGACCCGAATTTTCTGAAGT	2144
Db	1715	CCTACCAAGACACAGACTTCTCCGAAAT	1743

Search completed: September 6, 2004, 11:52:11
Job time : 762.842 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 6, 2004, 00:06:13 ; Search time 120.494 Seconds
(without alignments)
11537.137 Million cell updates/sec

Title: US-10-626-173-3

Perfect score: 2505

Sequence: 1 atgggaataataaactag.....gcagcttgaggacgcgcgc 2505

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/2/ina/6CTUS COMB.seq:*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2505	100.0	2505	3	US-09-291-839-3
2	2505	100.0	2505	4	US-09-458-457-3
3	2505	100.0	2505	4	US-09-947-199A-3
4	2505	100.0	3025	3	US-09-291-839-1
5	2505	100.0	3025	4	US-09-458-457-1
6	2505	100.0	3025	4	US-09-947-199A-1
7	1892.2	75.5	2505	4	US-09-458-457-9
8	1892.2	75.5	2505	4	US-09-947-199A-9
9	1892.2	75.5	3026	4	US-09-458-457-7
10	1892.2	75.5	3026	4	US-09-947-199A-7
11	427.4	17.1	616	4	US-09-833-381-1149
12	414.6	16.6	736	4	US-09-833-381-1150
13	413.2	16.5	740	4	US-09-833-381-1151
14	184.2	7.4	304	4	US-09-833-381-1147
15	84.4	3.4	186	4	US-09-833-381-1148
16	84	3.4	14770	4	US-09-220-132-30
17	78.8	3.1	2894	4	US-09-620-312D-95
18	78.8	3.1	3606	4	US-10-164-595-37
19	78.6	3.1	3454	3	US-09-082-059-1
20	75.4	3.0	4297	4	US-09-972-115A-3
21	75.4	3.0	4493	4	US-09-972-115A-1
22	74.2	3.0	384	4	US-09-372-115A-9
23	74.2	3.0	1069	4	US-09-833-381-841
24	74.2	3.0	3394	4	US-09-427-154-1
25	74.2	3.0	3498	4	US-09-350-982C-4
26	74.2	3.0	3797	4	US-09-696-668-1
27	74.2	3.0	3816	4	US-09-696-668-2

ALIGNMENTS

RESULT 1

US-09-291-839-3

; Sequence 3, Application US/09291839A

; Patent No. 6261818

; GENERAL INFORMATION:

; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-068

; CURRENT APPLICATION NUMBER: US/09/291,839A

; CURRENT FILING DATE: 1999-04-14

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 2505

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2505)

US-09-291-839-3

Query Match		100.0%;	Score 2505;	DB 3;	Length 2505;
Best Local Similarity		100.0%;	Pred No. 0;		
Matches 2505;		Conservative	0;	Mismatches	0; Gaps 0;
Qy	1	ATGGGAAATTATAAATCTTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAAGTC	60		
Db	1	ATGGGAAATTATAAATCTTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAAAGTC	60		
Qy	61	AGTGAATCATATGTTATACATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA	120		
Db	61	AGTGAATCATATGTTATACATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA	120		
Qy	121	GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATT	180		
Db	121	GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATT	180		
Qy	181	AATTACCGCACTGAATATGGCTGTCTTACTTCAATTTATGTTGATTTGGAGGCAAG	240		
Db	181	AATTACCGCACTGAATATGGCTGTCTTACTTCAATTTATGTTGATTTGGAGGCAAG	240		
Qy	241	AAATACATATTCGAACCTTATGTTGAAAGGGCTCCGCCCCTTCGACTGACAGAAT	300		
Db	241	AAATACATATTCGAACCTTATGTTGAAAGGGCTCCGCCCCTTCGACTGACAGAAT	300		
Qy	301	GGATTTACAGCCTTGCAATTTAGAGTAATGAGGTAATGAGGTAATGAGGTAATGAGG	360		
Db	301	GGATTTACAGCCTTGCAATTTAGAGTAATGAGGTAATGAGGTAATGAGGTAATGAGG	360		
Qy	361	CTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTCCTCCCTCATATT	420		

28	74.2	3.0	4275	4	US-09-972-115A-5	Sequence 5, Appli
29	74.2	3.0	4512	4	US-09-350-982C-3	Sequence 3, Appli
30	72.8	2.9	1056	3	US-09-065-474-138	Sequence 138, App
31	72.8	2.9	1056	3	US-09-065-474-140	Sequence 140, App
32	72.8	2.9	1056	4	US-09-557-034-138	Sequence 138, App
33	72.8	2.9	1056	4	US-09-557-034-140	Sequence 140, App
34	72.8	2.9	5235	1	US-09-031-485-35	Sequence 35, Appl
35	72.8	2.9	5235	1	US-09-031-485-36	Sequence 36, Appl
36	72.8	2.9	5235	1	US-08-847-429A-35	Sequence 35, Appl
37	72.8	2.9	5235	1	US-08-847-429A-36	Sequence 36, Appl
38	72.8	2.9	5235	3	US-09-065-474-35	Sequence 35, Appl
39	72.8	2.9	5235	3	US-09-065-474-36	Sequence 36, Appl
40	72.8	2.9	5235	4	US-09-557-034-35	Sequence 35, Appl
41	72.8	2.9	5235	4	US-09-557-034-36	Sequence 36, Appl
42	72.8	2.9	5503	1	US-09-031-485-32	Sequence 32, Appl
43	72.8	2.9	5503	1	US-09-031-485-34	Sequence 34, Appl
44	72.8	2.9	5503	1	US-08-847-429A-32	Sequence 32, Appl
45	72.8	2.9	5503	1	US-08-847-429A-34	Sequence 34, Appl

361	Db	CTTCAAGTGGAGCTGATATACAGCAGGTGGATACGGTGGCCCTCACTGGCCCTCCATATT	420	1441	CGATGCAGAAATAAAATAGTGGCTATATAAACGGTTATGAGGCCAATACCTTACTGCTCCAAG	1500
421	Qy	GCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGGCAACATGGAAGCTAAATGTC	480	1501	TCAGATGTGGATATGTTTTGGCCGAGAGGTGTCCAATCTCTGCCAGCTCAATCATCTCCCTGC	1560
421	Db	GCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGGCAACATGGAAGCTAAATGTC	480	1501	TCAGATGTGGATATGTTTTGGCCGAGAGGTGTCCAATCTCTGCCAGCTCAATCATCTCCCTGC	1560
481	Qy	AATATTCAAGATGCAGTGTGTTTTTCACTCCATTGCATATTGACGGCTCATATGGACATGAA	540	1561	GTAATTGATGTTGGGTGCTTGTGAAATGATCCAGCCAGTTTGGCAATGTCTCACTCAA	1620
481	Db	AATATTCAAGATGCAGTGTGTTTTTCACTCCATTGCATATTGACGGCTCATATGGACATGAA	540	1561	GTAATTGATGTTGGGTGCTTGTGAAATGATCCAGCCAGTTTGGCAATGTCTCACTCAA	1620
541	Qy	CAGGTAACCTGCGCTTCTTTTGAAATTTGGTGTGATGTAATGTAAGTGAAGTGGAAAGTTGGA	600	1621	TACATATCAGGGGGTCTCTGTTCTCCCTCTTTCATGAGCAGAAAGAGGATTTCTTGATTTG	1680
541	Db	CAGGTAACCTGCGCTTCTTTTGAAATTTGGTGTGATGTAATGTAAGTGAAGTGGAAAGTTGGA	600	1621	TACATATCAGGGGGTCTCTGTTCTCCCTCTTTCATGAGCAGAAAGAGGATTTCTTGATTTG	1680
601	Qy	GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAAACTCTTGATG	660	1681	CAGTCTAAATTAATTATGCAGTAGATGTTGGCCAAAGGATGAGTACTTCCAAACCTG	1740
601	Db	GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGCAAAACTCTTGATG	660	1681	CAGTCTAAATTAATTATGCAGTAGATGTTGGCCAAAGGATGAGTACTTCCAAACCTG	1740
661	Qy	GAAGAAAGGAGCAAAAGCATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT	720	1741	ACAGACCAATATACATCGTTCGAAACAGTCACAATATTCTTCTATGAGGATGGG	1800
661	Db	GAAGAAAGGAGCAAAAGCATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT	720	1741	ACAGACCAATATACATCGTTCGAAACAGTCACAATATTCTTCTATGAGGATGGG	1800
721	Qy	TTCTGTCTTCGATTTGGACACCATATAGTTAAAGTATCTGCTGCAAGGATTTGGAA	780	1801	CATGCTGTGTGGCAGATTTTGGAGAAATTAAGATTTCTACAGTCTCTGGATGAAGACAAC	1860
721	Db	TTCTGTCTTCGATTTGGACACCATATAGTTAAAGTATCTGCTGCAAGGATTTGGAA	780	1801	CATGCTGTGTGGCAGATTTTGGAGAAATTAAGATTTCTACAGTCTCTGGATGAAGACAAC	1860
781	Qy	GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGTCAAT	840	1861	ATGACAAAACAACCTGSGAACCTCCGTGGATGGCTCTGAGGTGTTTCACGAGTGCACCT	1920
781	Db	GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGTCAAT	840	1861	ATGACAAAACAACCTGSGAACCTCCGTGGATGGCTCTGAGGTGTTTCACGAGTGCACCT	1920
841	Qy	GGCAAAATTTGAAGTTCGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG	900	1921	CGGTACACCATCAAAAGCAGATGCTTTCAGCTATGCTCTGTGTCTGTGGGAAATTTCTCACT	1980
841	Db	GGCAAAATTTGAAGTTCGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG	900	1921	CGGTACACCATCAAAAGCAGATGCTTTCAGCTATGCTCTGTGTCTGTGGGAAATTTCTCACT	1980
901	Qy	GAAAACATCTTCAGTGAACACAGCTTTTTCATAGTGTCTGTACTATGGCAAGCAATGAC	960	1981	GCGGAAATTCATTTGCGTCAATCAAGCCAGCGGCTGCGGCAGCAGACATGSGTTACCAAC	2040
901	Db	GAAAACATCTTCAGTGAACACAGCTTTTTCATAGTGTCTGTACTATGGCAAGCAATGAC	960	1981	GCGGAAATTCATTTGCGTCAATCAAGCCAGCGGCTGCGGCAGCAGACATGSGTTACCAAC	2040
961	Qy	CTAGTCAAAATTTCTTGTGATCAGAAATGTCATAAATCATCAACCAAGGAGGATGGG	1020	2041	CACATCAGACCTCCCATTTGGCTATTTCATTTCCAAAGCCCATATCATCTCTGTGTATACGA	2100
961	Db	CTAGTCAAAATTTCTTGTGATCAGAAATGTCATAAATCATCAACCAAGGAGGATGGG	1020	2041	CACATCAGACCTCCCATTTGGCTATTTCATTTCCAAAGCCCATATCATCTCTGTGTATACGA	2100
1021	Qy	CACACTGGATTACACTCTGCTTGTACACAGGTTCATTTGCGCTGGTTCACTTCTACTG	1080	2101	GCGTGAACCGCATGTCCTGAAGGAAGACCCGAAATTTTCTGAAAGTTGTCTATGAAGTTAGAA	2160
1021	Db	CACACTGGATTACACTCTGCTTGTACACAGGTTCATTTGCGCTGGTTCACTTCTACTG	1080	2101	GCGTGAACCGCATGTCCTGAAGGAAGACCCGAAATTTTCTGAAAGTTGTCTATGAAGTTAGAA	2160
1081	Qy	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATGCCAGCAGGCTTAGTGGTGA AAAA	1140	2161	GAGTGTCTCTGCAACATTTGAGCTGATGTCCTCTGATCAAGTAAACAGCAGTGGTCTCTC	2220
1081	Db	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATGCCAGCAGGCTTAGTGGTGA AAAA	1140	2161	GAGTGTCTCTGCAACATTTGAGCTGATGTCCTCTGATCAAGTAAACAGCAGTGGTCTCTC	2220
1141	Qy	GATGACGAGACATGTTGATGTGGCTTATGAAAAGGCAATGATGCCATTGTCACACTC	1200	2221	TCACCTTCTTCTTCTTCTGATTGCTTGGTGAACCGGGGAGGACCTGGCCGGAGTCAATGG	2280
1141	Db	GATGACGAGACATGTTGATGTGGCTTATGAAAAGGCAATGATGCCATTGTCACACTC	1200	2221	TCACCTTCTTCTTCTTCTGATTGCTTGGTGAACCGGGGAGGACCTGGCCGGAGTCAATGG	2280
1201	Qy	CTGAAGCATTTATAAGAGACCAAGATGAATTTGCCCTGTATGAATATTCTCAGCCTGGA	1260	2281	GCAGCATTTAAGAGTGGTTCGAAATTTGCAATATGCTCTAAATGCAAGTCCCTATGCTGCT	2340
1201	Db	CTGAAGCATTTATAAGAGACCAAGATGAATTTGCCCTGTATGAATATTCTCAGCCTGGA	1260	2281	GCAGCATTTAAGAGTGGTTCGAAATTTGCAATATGCTCTAAATGCAAGTCCCTATGCTGCT	2340
1261	Qy	GAGAGTGGCTCTATGTGCTGTTCATCACCCTTGGGAGATTAAGCATGACAAAA	1320	2341	TTGTCCCAAAGTGTGGAACAATATTCCTCTCAAGGTCTGTCTTTTGGAGGAGATGAAAAGA	2400
1261	Db	GAGAGTGGCTCTATGTGCTGTTCATCACCCTTGGGAGATTAAGCATGACAAAA	1320	2341	TTGTCCCAAAGTGTGGAACAATATTCCTCTCAAGGTCTGTCTTTTGGAGGAGATGAAAAGA	2400
1321	Qy	GAGAAGGCAGATATTCTCCTCTCAAGAGCTGGAATTCGCTTCACTTTCCATCTTCAGCTC	1380	2401	AGTCTTCAATACACCCCATTTGACAAAATATGGCTATGATCCGATCCCATGAGCTCAATG	2460
1321	Db	GAGAAGGCAGATATTCTCCTCTCAAGAGCTGGAATTCGCTTCACTTTCCATCTTCAGCTC	1380	2401	AGTCTTCAATACACCCCATTTGACAAAATATGGCTATGATCCGATCCCATGAGCTCAATG	2460
1381	Qy	TCAGAAATTTGAGTTCCATGAGATATTGCTCAGGTTCTTTTGGGAAAGTATATAAGGA	1440	2461	CATTTTCAATCTTCCGAAAATAGTAGCAGCTTTTGAAGCAGCAGCAGC 2505	
1381	Db	TCAGAAATTTGAGTTCCATGAGATATTGCTCAGGTTCTTTTGGGAAAGTATATAAGGA	1440	2461	CATTTT	

```
; Sequence 3, Application US/09458457
; Patent No. 6500654
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL DARK PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: mni-068cp
; CURRENT APPLICATION NUMBER: US/09/458,457
; EARLIER FILING DATE: 1999-12-10
; EARLIER APPLICATION NUMBER: 60/111,938
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/291,839
; EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2505)
US-09-458-457-3

Query Match      100.0%; Score 2505; DB 4; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGGGAATTTATAAATCTAGACCAACCCAAACTTGTACTGATGAATGGAGAAAAAGTC 60
DB      1  ATGGGAATTTATAAATCTAGACCAACCCAAACTTGTACTGATGAATGGAGAAAAAGTC 60

QY      61  AGTGAATCATATGTTATCACAATAGAAAAGATTAGAAAGTGAACCTGCGAGATCAAGAAAAA 120
DB      61  AGTGAATCATATGTTATCACAATAGAAAAGATTAGAAAGTGAACCTGCGAGATCAAGAAAAA 120

QY      121  GAATGACAGAACTAAGGATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA 180
DB      121  GAATGACAGAACTAAGGATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTTA 180

QY      181  AATTACCGCACTGAAATGGGCTGTCTACTCTTCACTTTCACTTTATGTTGCAATTTT 240
DB      181  AATTACCGCACTGAAATGGGCTGTCTACTCTTCACTTTCACTTTATGTTGCAATTTT 240

QY      241  AAATCAATATTCGAATCTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAAT 300
DB      241  AAATCAATATTCGAATCTTATGTTGAAAGGGCTCCGCCATCTCGACTGACAGAAAT 300

QY      301  GGATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCAGAAATTGATCACTTCTCTG 360
DB      301  GGATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCAGAAATTGATCACTTCTCTG 360

QY      361  CTTACAGTGGAGCTGATATACAGCAGTTTGGATACGGTGGCTCACTGCCCTCCATATT 420
DB      361  CTTACAGTGGAGCTGATATACAGCAGTTTGGATACGGTGGCTCACTGCCCTCCATATT 420

QY      421  GCTACAAATAGCTGCCACCTAGAGGCTGTGATGCTGTGTCACATGGAGCTTAATGTC 480
DB      421  GCTACAAATAGCTGCCACCTAGAGGCTGTGATGCTGTGTCACATGGAGCTTAATGTC 480

QY      481  AATATTCAAGATCAGTTTTTTTTTCACTCCATGTCATATTGACGGTACTATGGACATGAA 540
DB      481  AATATTCAAGATCAGTTTTTTTTTCACTCCATGTCATATTGACGGTACTATGGACATGAA 540

QY      541  CAGGTAATCGCTTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGA 600
DB      541  CAGGTAATCGCTTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGA 600

QY      601  GATAGACCCCTCCACCTAGCATCTGCAAAAGAGTCTTGAATATTGCAAAAATCTTTGATG 660
DB      601  GATAGACCCCTCCACCTAGCATCTGCAAAAGAGTCTTGAATATTGCAAAAATCTTTGATG 660

QY      661  GAAGAAGGCAGAAAGCAGATGTAATGCTCAAGATAATGAAGACCAATGTCCTCCACTCCAT 720
```

```
DB      661  GAAGAAGGCAGAAAGCAGATGTAATGCTCAAGATAATGAAGACCAATGTCCTCCACTCCAT 720
QY      721  TTCTGTCTTCGATTTGGACACCATGATATAGTTAAGTATCTCTGCTGCAAAAGTGAATTTG 780
DB      721  TTCTGTCTTCGATTTGGACACCATGATATAGTTAAGTATCTCTGCTGCAAAAGTGAATTTG 780
QY      781  GTTCAAACTCATCTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAAT 840
DB      781  GTTCAAACTCATCTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACAAAT 840
QY      841  GGCAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
DB      841  GGCAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
QY      901  GAAAAACATCTTCAGTGAACACAGCTTTTCACTAGTGGCTGTGATGCTATGGCAAGAGCATGAC 960
DB      901  GAAAAACATCTTCAGTGAACACAGCTTTTCACTAGTGGCTGTGATGCTATGGCAAGAGCATGAC 960
QY      961  CTAGTCAAAATTTCTTCTTGTATCAGAAATGTCAATAACATCAACCAAGGAAGGATGGG 1020
DB      961  CTAGTCAAAATTTCTTCTTGTATCAGAAATGTCAATAACATCAACCAAGGAAGGATGGG 1020
QY      1021  CACACTGGATTACACTCTGCTTGTACACGGTCAATTCGGCTGGTTTCAGTCTTACTG 1080
DB      1021  CACACTGGATTACACTCTGCTTGTACACGGTCAATTCGGCTGGTTTCAGTCTTACTG 1080
QY      1081  GATAATGGAGCTCATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAATAA 1140
DB      1081  GATAATGGAGCTCATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAATAA 1140
QY      1141  GATGAGCAGACATGTTGATGTTGGCTTATGAAAAAGGSCATGATGCCATTTGCACATC 1200
DB      1141  GATGAGCAGACATGTTGATGTTGGCTTATGAAAAAGGSCATGATGCCATTTGCACATC 1200
QY      1201  CTGAAGCATTATTAAGAGACCAAGAGATGAATTCCTGCTGTAATGAATATTTCTCAGCTCGA 1260
DB      1201  CTGAAGCATTATTAAGAGACCAAGAGATGAATTCCTGCTGTAATGAATATTTCTCAGCTCGA 1260
QY      1261  GGAGATGGCTCTTATGTTGCTTCCATCACCTTGGGGAAGATTAAAGCAGATGACAAAA 1320
DB      1261  GGAGATGGCTCTTATGTTGCTTCCATCACCTTGGGGAAGATTAAAGCAGATGACAAAA 1320
QY      1321  GAGAAGCAGATATTCCTCTTACAGCTGATGCTTCCATTTCCATTTCCATTTCCATTTCC 1380
DB      1321  GAGAAGCAGATATTCCTCTTACAGCTGATGCTTCCATTTCCATTTCCATTTCCATTTCC 1380
QY      1381  TCAGAAATTTGAGTTCCATGAGATTTATGGCTCAGGTTCCTTTGGGAAAGTATATAAGGA 1440
DB      1381  TCAGAAATTTGAGTTCCATGAGATTTATGGCTCAGGTTCCTTTGGGAAAGTATATAAGGA 1440
QY      1441  CGATGAGAAATTAATAATAGTGGCTATAAAACGTTATCGAGCCCAATACCTACTGCTCCAAG 1500
DB      1441  CGATGAGAAATTAATAATAGTGGCTATAAAACGTTATCGAGCCCAATACCTACTGCTCCAAG 1500
QY      1501  TCAGATGTGATATGTTTGGCGAGAGGTGTCATTTCTGCGAGCTCAATCATCCCTGC 1560
DB      1501  TCAGATGTGATATGTTTGGCGAGAGGTGTCATTTCTGCGAGCTCAATCATCCCTGC 1560
QY      1561  GTAATTCAAGTTGTGGGTGCTTGTGAATGATCCAGCCAGTTGGCCATGTTGCTACTCAA 1620
DB      1561  GTAATTCAAGTTGTGGGTGCTTGTGAATGATCCAGCCAGTTGGCCATGTTGCTACTCAA 1620
QY      1621  TACATATCAGGGGGTCTCTGTTCTCCCTTCTATGAGCAGAGAGGATTTCTGATTTG 1680
DB      1621  TACATATCAGGGGGTCTCTGTTCTCCCTTCTATGAGCAGAGAGGATTTCTGATTTG 1680
QY      1681  CAGTCTAAATTAATTTAGTCAAGTAGATGTTGCCAAAGCATGGAGTACTTTCAACAACCTG 1740
DB      1681  CAGTCTAAATTAATTTAGTCAAGTAGATGTTGCCAAAGCATGGAGTACTTTCAACAACCTG 1740
QY      1741  ACACAGCCATTTATACATGCTGACTTGAACAGTCAATATTTCTTCTATGAGAGATGGG 1800
```

Db 1741 ACACAGCCAAATATACATCGTGACCTTGAACAGTCACAATATCTTCTCTATGAGGATGG 1800
QY 1801 CATGCTGGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
Db 1801 CATGCTGGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
QY 1861 ATGACAAAACCACTGGAAACCTCGTTGGATGCTCTGAGGTGTTTCAGCAGTGCACCT 1920
Db 1861 ATGACAAAACCACTGGAAACCTCGTTGGATGCTCTGAGGTGTTTCAGCAGTGCACCT 1920
QY 1921 CGGTACACCAATCAAAAGCAGATGTCTTTCAGCTATGCTCTGTGTCTGTGGGAAATTCACCT 1980
Db 1921 CGGTACACCAATCAAAAGCAGATGTCTTTCAGCTATGCTCTGTGTCTGTGGGAAATTCACCT 1980
QY 1981 GGGGAAATTCATTCGCTCATCTCAAGCCAGGGCTGGCAGCAGACATGGCTTACCAC 2040
Db 1981 GGGGAAATTCATTCGCTCATCTCAAGCCAGGGCTGGCAGCAGACATGGCTTACCAC 2040
QY 2041 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGTGATACGA 2100
Db 2041 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCATATCATCTCTGTGATACGA 2100
QY 2101 GGGTGAACGCAATGCTTGAAGAGACCCGAATTTCTGAAATGTCATGAAGTTAGAA 2160
Db 2101 GGGTGAACGCAATGCTTGAAGAGACCCGAATTTCTGAAATGTCATGAAGTTAGAA 2160
QY 2161 GAGTGTCTCTGCAACATTTGAGTGTCTCTGATGCTCTGATCAAGTAAACAGCAGTGGTCTCTC 2220
Db 2161 GAGTGTCTCTGCAACATTTGAGTGTCTCTGATGCTCTGATCAAGTAAACAGCAGTGGTCTCTC 2220
QY 2221 TCACCTTCTTCTTCTGATGCTGTGTAACCGGGAGGACCTGCGCGGAGTCATGTG 2280
Db 2221 TCACCTTCTTCTTCTGATGCTGTGTAACCGGGAGGACCTGCGCGGAGTCATGTG 2280
QY 2281 GCAGCATTAAGAGTCGTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTGT 2340
Db 2281 GCAGCATTAAGAGTCGTTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTGT 2340
QY 2341 TTGTCCCAAGTGTGGAACAATATTCCTCTCAAGTCTGTCTTTTGGAGGAGTGAAGA 2400
Db 2341 TTGTCCCAAGTGTGGAACAATATTCCTCTCAAGTCTGTCTTTTGGAGGAGTGAAGA 2400
QY 2401 AGTCTTCAATACACACCCATTCGAATATGCTATGCTATGCTATCCGATCCATGAGTCAATG 2460
Db 2401 AGTCTTCAATACACACCCATTCGAATATGCTATGCTATGCTATCCGATCCATGAGTCAATG 2460
QY 2461 CATTTTTCATTTCTTGGCGGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505
Db 2461 CATTTTTCATTTCTTGGCGGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505

RESULT 3

US-09-947-199A-3
; Sequence 3, Application US/09947199A
; Patent No. 6660490
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199A
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199A-3

Query Match 100.0%; Score 2505; DB 4; Length 2505;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAATTTATAATCTAGACCAACCCAAACTTGTACTGATGAGTGGAGAAAAGTC 60
Db 1 ATGGGAAATTTATAAATCTAGACCAACCCAAACTTGTACTGATGAGTGGAGAAAAGTC 60
QY 61 AGTCAATCATATGTTATCACAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 120
Db 61 AGTCAATCATATGTTATCACAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 120
QY 121 GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCACTAAAGTCAATT 180
Db 121 GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCACTAAAGTCAATT 180
QY 181 AATTACCGCACTGAAATGGCTGTCTTACTTCTTATTTGTCATTTTGTGGAGGCAAG 240
Db 181 AATTACCGCACTGAAATGGCTGTCTTACTTCTTATTTGTCATTTTGTGGAGGCAAG 240
QY 241 AAATCACATATTCGAACCTCTTATTTGAAAGGGCTCCGCCATCTCGACTGACAAGAA 300
Db 241 AAATCACATATTCGAACCTCTTATTTGAAAGGGCTCCGCCATCTCGACTGACAAGAA 300
QY 301 GGATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCAAGATTGATGATCACTTCTG 360
Db 301 GGATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCAAGATTGATGATCACTTCTG 360
QY 361 CTTTCAAGTGGAGCTGATATACAGCAGGTTGAGTACGGTGGCTCACTGCCCTCCATATT 420
Db 361 CTTTCAAGTGGAGCTGATATACAGCAGGTTGAGTACGGTGGCTCACTGCCCTCCATATT 420
QY 421 GCTCAATAGTGGCCACCTAGAGGCTGCTGATGCTGTGCAACATGGAGTCAATGTC 480
Db 421 GCTCAATAGTGGCCACCTAGAGGCTGCTGATGCTGTGCAACATGGAGTCAATGTC 480
QY 481 AATATTCAAGATGCAAGTCTTTTCACTCCATTCATATTGCAGCGTACTATGACATGAA 540
Db 481 AATATTCAAGATGCAAGTCTTTTCACTCCATTCATATTGCAGCGTACTATGACATGAA 540
QY 541 CAGTAACTCGCCTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGA 600
Db 541 CAGTAACTCGCCTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGA 600
QY 601 GATAGCCCTCCACCTAGCATCTGCAAGAGGATTCTTGAATTTCCAAACTCTTGATG 660
Db 601 GATAGCCCTCCACCTAGCATCTGCAAGAGGATTCTTGAATTTCCAAACTCTTGATG 660
QY 661 GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720
Db 661 GAAGAAGCAGCAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720
QY 721 TTCTGTTCTCGATTGGACACCATTATAGTTAAGTATCTGCTGCAAGATGTTGGA 780
Db 721 TTCTGTTCTCGATTGGACACCATTATAGTTAAGTATCTGCTGCAAGATGTTGGA 780
QY 781 GTTCAACTCATGTTGTTAATATCTATGGAGATACCCCTTACACTGGCAGTCTACAT 840
Db 781 GTTCAACTCATGTTGTTAATATCTATGGAGATACCCCTTACACTGGCAGTCTACAT 840
QY 841 GGCAGAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGACGAAAGTCTGACTAG 900
Db 841 GGCAGAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGACGAAAGTCTGACTAG 900
QY 901 GAAACCATCTTCAGTGAACAGCTTTTTCATAGTCTTTGTACCTATGGCAGAGCATTTGAC 960
Db 901 GAAACCATCTTCAGTGAACAGCTTTTTCATAGTCTTTGTACCTATGGCAGAGCATTTGAC 960

QY 961 CTAGTCAAAATTTCTTTGATCAGAAATGTCATTAACATCAACACCAAGGAAGGATGG 1020
Db 961 CTAGTCAAAATTTCTTTGATCAGAAATGTCATTAACATCAACACCAAGGAAGGATGG 1020
QY 1021 CACACTGATTAACACTCTGCTTGTACCAAGGTGCACATTCGCTGTTTCACTTCTTACTG 1080
Db 1021 CACACTGATTAACACTCTGCTTGTACCAAGGTGCACATTCGCTGTTTCACTTCTTACTG 1080
QY 1081 GATATGAGCTGATATGAATCTAGTGGCTTATGAAAGGCGCATGATGCCATTGTCACTC 1140
Db 1081 GATATGAGCTGATATGAATCTAGTGGCTTATGAAAGGCGCATGATGCCATTGTCACTC 1140
QY 1141 GATGAGCAGACATGTTTGTATGAGCTTATGAAAGGCGCATGATGCCATTGTCACTC 1200
Db 1141 GATGAGCAGACATGTTTGTATGAGCTTATGAAAGGCGCATGATGCCATTGTCACTC 1200
QY 1201 CTGAAGCATTAAGAGACCAAGATGATGTCCTGTTTGAATGAAATTCCTGAGCTGGA 1260
Db 1201 CTGAAGCATTAAGAGACCAAGATGATGTCCTGTTTGAATGAAATTCCTGAGCTGGA 1260
QY 1261 GGAGATGGCTCTATGCTGCTTCCATCACCTTGGGGAAGATTAAAGCATGACAAA 1320
Db 1261 GGAGATGGCTCTATGCTGCTTCCATCACCTTGGGGAAGATTAAAGCATGACAAA 1320
QY 1321 GAGAAGCAGATATTCCTCTTAAGAGCTGGATGCTTCACTTCCATCTTCACTC 1380
Db 1321 GAGAAGCAGATATTCCTCTTAAGAGCTGGATGCTTCACTTCCATCTTCACTC 1380
QY 1381 TCAGAAATGAGTCCATGAGATTAATGCTGCTGAGTTCCTTGGGAAAGTATATAAGGA 1440
Db 1381 TCAGAAATGAGTCCATGAGATTAATGCTGCTGAGTTCCTTGGGAAAGTATATAAGGA 1440
QY 1441 CGATGCAAAATAAATAGTGGCTTATAAAGCTTATCGAGCCCAATACCTACTGCTCCAA 1500
Db 1441 CGATGCAAAATAAATAGTGGCTTATAAAGCTTATCGAGCCCAATACCTACTGCTCCAA 1500
QY 1501 TCAGATGATATGTTTTCGCGAGAGTGTCAATCTCTGCGAGCTCAATCACTCCCTG 1560
Db 1501 TCAGATGATATGTTTTCGCGAGAGTGTCAATCTCTGCGAGCTCAATCACTCCCTG 1560
QY 1561 GTAATTCAGTTTGGGTGCTTGTCAATGATCCAGCAGTTTCCATTTGTCCTCAA 1620
Db 1561 GTAATTCAGTTTGGGTGCTTGTCAATGATCCAGCAGTTTCCATTTGTCCTCAA 1620
QY 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGAGCAGAGGATTTCTGATTG 1680
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGAGCAGAGGATTTCTGATTG 1680
QY 1681 CAGTCTAAATTAATATTGAGTAGATGTTGCCAAGGAGTGGATACCTTCAACACCTG 1740
Db 1681 CAGTCTAAATTAATATTGAGTAGATGTTGCCAAGGAGTGGATACCTTCAACACCTG 1740
QY 1741 ACACAGCAATTAATACATCGTACCTGAAACAGTCAATATTTCTCTATGAGGATGG 1800
Db 1741 ACACAGCAATTAATACATCGTACCTGAAACAGTCAATATTTCTCTATGAGGATGG 1800
QY 1801 CATCTGTGTGGGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAGAACAC 1860
Db 1801 CATCTGTGTGGGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAGAACAC 1860
QY 1861 ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTCAAGGATGCACT 1920
Db 1861 ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTCAAGGATGCACT 1920
QY 1921 CGGTACACCAATCAAGCAGATGCTTCACTATGCTGCTGCTGTTGGGAAATTTCTCACT 1980
Db 1921 CGGTACACCAATCAAGCAGATGCTTCACTATGCTGCTGCTGTTGGGAAATTTCTCACT 1980
QY 1981 GGGCAATTCATTCGCTCATCTCAAGCAGGCTGGGACGACACATGCTTACAC 2040
Db 1981 GGGCAATTCATTCGCTCATCTCAAGCAGGCTGGGACGACACATGCTTACAC 2040

QY 2041 CACATCAGACCTCCCATTTGGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATAACA 2100
Db 2041 CACATCAGACCTCCCATTTGGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATAACA 2100
QY 2101 GGGTGGAAACGATCTCTGAAAGAACCCGAAATTTTCTGAAGTTGTCATGAAGTTAGAA 2160
Db 2101 GGGTGGAAACGATCTCTGAAAGAACCCGAAATTTTCTGAAGTTGTCATGAAGTTAGAA 2160
QY 2161 GAGTGTCTCTGACATGAGCTGATGTCCTGCAATCAAGTAAACAGCAGTGGGTCTCTC 2220
Db 2161 GAGTGTCTCTGACATGAGCTGATGTCCTGCAATCAAGTAAACAGCAGTGGGTCTCTC 2220
QY 2221 TCACCTTCTTCTTCTGATTTGCTGAAACCGGGGAGACCTGGCCGAGTCAATG 2280
Db 2221 TCACCTTCTTCTTCTGATTTGCTGAAACCGGGGAGACCTGGCCGAGTCAATG 2280
QY 2281 GCACATTAAGAGTCTGTTTTCGAATTCGAATATGCTCTTAAATGCAAGGTCCTATGCTCT 2340
Db 2281 GCACATTAAGAGTCTGTTTTCGAATTCGAATATGCTCTTAAATGCAAGGTCCTATGCTCT 2340
QY 2341 TTGTCCCAAAAGTGTGCAATATTCCTCTCAAGGTCCTGTTTGGAGGAGATGAAAGA 2400
Db 2341 TTGTCCCAAAAGTGTGCAATATTCCTCTCAAGGTCCTGTTTGGAGGAGATGAAAGA 2400
QY 2401 AGTCTTCAATACACACCATTTGAATAATGCTATGCTATCCGATCCCATGAGCTCAATG 2460
Db 2401 AGTCTTCAATACACACCATTTGAATAATGCTATGCTATCCGATCCCATGAGCTCAATG 2460
QY 2461 CATTTTCATTTCTGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505
Db 2461 CATTTTCATTTCTGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505

RESULT 4

US-09-291-839-1

; Sequence 1, Application US/09291839A

; Patent No. 6261818

; GENERAL INFORMATION:

; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: WNI-068

; CURRENT APPLICATION NUMBER: US/09/291,839A

; CURRENT FILING DATE: 1999-04-14

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 3025

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (48)..(2552)

US-09-291-839-1

Query Match 100.0%; Score 2505; DB 3; Length 3025;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAATTTATAATCTAGACCAACCCAACTTGTATGATGAATGGAAGAAAAAGTC 60

Db 48 ATGGGAATTTATAATCTAGACCAACCCAACTTGTATGATGAATGGAAGAAAAAGTC 107

QY 61 AGTGAATCATCTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120

Db 108 AGTGAATCATCTTATCAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 167

QY 121 GAATGACAGAACTAAGGAATATATTGGCTCTGTATGAAGCCTTCAGTAAAGTCAATTTA 180

Db 168 GAATGACAGAACTAAGGAATATATTGGCTCTGTATGAAGCCTTCAGTAAAGTCAATTTA 227

QY 181 AATTACCGCACTGAAAATGSGCTGTCTTACTTCAATTATGTTGCAATTTGTGGAGGCAAG 240

Db 228 AATTACCGCACTGAAAATGSGCTGTCTTACTTCAATTATGTTGCAATTTGTGGAGGCAAG 287

QY	241	AAATCACATATTCGAATCTCTTATGTTGAAAGGGCTCGGCCCATCTCGACTGACAGAAAT	300	QY	1321	GAGAAGGCGAGATATTCCTCTCTTAAGAGCTGGATTCGCTTCAATTTCCATCTTCAGCTC	1380
Db	288	AAATCACATATTCGAATCTCTTATGTTGAAAGGGCTCGGCCCATCTCGACTGACAGAAAT	347	Db	1368	GAGAAGGCGAGATATTCCTCTCTTAAGAGCTGGATTCGCTTCAATTTCCATCTTCAGCTC	1427
QY	301	GGATTACAGCCTTCGATTTAGCAGTTTACAAGGATAATGCGAGAAATGATCTCTCTG	360	QY	1381	TCAGAAATTCAGTTTCCATGAGATATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA	1440
Db	348	GGATTACAGCCTTCGATTTAGCAGTTTACAAGGATAATGCGAGAAATGATCTCTCTG	407	Db	1428	TCAGAAATTCAGTTTCCATGAGATATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA	1487
QY	361	CTTCACATGGAGCTGATATACAGCAGGTTGGATACGGTGGCTTCATCTGCCCTCCATAT	420	QY	1441	CGATGCGAAATAAAATAGTGGCTATATAAAAGTTATCGAGCCCAATACCTACTCTCCAAAG	1500
Db	408	CTTCACATGGAGCTGATATACAGCAGGTTGGATACGGTGGCTTCATCTGCCCTCCATAT	467	Db	1488	CGATGCGAAATAAAATAGTGGCTATATAAAAGTTATCGAGCCCAATACCTACTCTCCAAAG	1547
QY	421	GCTACAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTTGCAACATGAGAGCTAATGTC	480	QY	1501	TCAGATGTGATATGTTTTCGCGAGAGTCTCCATCTCTGCGAGCTCAATCATCCCTGC	1560
Db	468	GCTACAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTTGCAACATGAGAGCTAATGTC	527	Db	1548	TCAGATGTGATATGTTTTCGCGAGAGTCTCCATCTCTGCGAGCTCAATCATCCCTGC	1607
QY	481	AATATTCAGATGCGAGTTTCTTCTGCTGATGTAATGTAAGTGGTGAAGTTGGA	540	QY	1561	GTAATTCAGTTTGGGTGCTTGTGTAATGATCCAGCCAGTTTGCCTACTCAAA	1620
Db	528	AATATTCAGATGCGAGTTTCTTCTGCTGATGTAATGTAAGTGGTGAAGTTGGA	587	Db	1608	GTAATTCAGTTTGGGTGCTTGTGTAATGATCCAGCCAGTTTGCCTACTCAAA	1667
QY	541	CAGGTAACCTCGCTTCTTCTGCTGATGTAATGTAAGTGGTGAAGTTGGA	600	QY	1621	TACATATCAGGGGTTCTCTGTTCTCCCTCATGAGCAGAGAGGATTTCTGATTTG	1680
Db	588	CAGGTAACCTCGCTTCTTCTGCTGATGTAATGTAAGTGGTGAAGTTGGA	647	Db	1668	TACATATCAGGGGTTCTCTGTTCTCCCTCATGAGCAGAGAGGATTTCTGATTTG	1727
QY	601	GATAGACCCCTCCACTAGCATCTGCAAAAGGATCTTGAATATGTCAAACTCTTGATG	660	QY	1681	CAGTCTAAATTAATATTCAGTAGATGTTGCCAAAGGATGGAGTACCTTCAACACCTG	1740
Db	648	GATAGACCCCTCCACTAGCATCTGCAAAAGGATCTTGAATATGTCAAACTCTTGATG	707	Db	1728	CAGTCTAAATTAATATTCAGTAGATGTTGCCAAAGGATGGAGTACCTTCAACACCTG	1787
QY	661	GAAGAGGCGACAGACGATGTAATGCTCAAGATATGAAGACCATGTCCTCCACTCCAT	720	QY	1741	ACACAGCCAAATPACATCGTAGCTTGAAACAGTCAAAATATTTCTTCTATGAGGATGGG	1800
Db	708	GAAGAGGCGACAGACGATGTAATGCTCAAGATATGAAGACCATGTCCTCCACTCCAT	767	Db	1788	ACACAGCCAAATPACATCGTAGCTTGAAACAGTCAAAATATTTCTTCTATGAGGATGGG	1847
QY	721	TTCTGTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTTGAA	780	QY	1801	CATCGTGTGGGCGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAAC	1860
Db	768	TTCTGTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTTGAA	827	Db	1848	CATCGTGTGGGCGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGCAAC	1907
QY	781	GTTCACCTCATGTTGTTAATATCTATGAGATACCCCTTACACCTGGCATGTACAAT	840	QY	1861	ATGCAAAACCAACCTGGGAACTCCGTTGGATGGCTCTGAGGTGTCAACGAGTGCAT	1920
Db	828	GTTCACCTCATGTTGTTAATATCTATGAGATACCCCTTACACCTGGCATGTACAAT	887	Db	1908	ATGCAAAACCAACCTGGGAACTCCGTTGGATGGCTCTGAGGTGTCAACGAGTGCAT	1967
QY	841	GGCAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAAACAGAAAGTCTCACTAAG	900	QY	1921	CGGTACACCATCAAGCAGATGTTCTCAGCTATGCTTGTGTCTGTGGGAAATTTCTCACT	1980
Db	888	GGCAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAAACAGAAAGTCTCACTAAG	947	Db	1968	CGGTACACCATCAAGCAGATGTTCTCAGCTATGCTTGTGTCTGTGGGAAATTTCTCACT	2027
QY	901	GAAAACATCTTCAGTGAAACAGCTTTTCATAGTGTGTAATGCTATGCGCAGAGCATTTGAC	960	QY	1981	GGCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGCGCAGCAGCATGGCTTACCAC	2040
Db	948	GAAAACATCTTCAGTGAAACAGCTTTTCATAGTGTGTAATGCTATGCGCAGAGCATTTGAC	1007	Db	2028	GGCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTGGCGCAGCAGCATGGCTTACCAC	2087
QY	961	CTAGTCAAAATTTCTTCTGATCAGAAATGTCTATAACATCAACCAAGGAGGATGGG	1020	QY	2041	CACATCAGACCTCCCATTTGGCTATTTCCATTTCCAGCCCATATCATCTCTGCTGATACGA	2100
Db	1008	CTAGTCAAAATTTCTTCTGATCAGAAATGTCTATAACATCAACCAAGGAGGATGGG	1067	Db	2088	CACATCAGACCTCCCATTTGGCTATTTCCATTTCCAGCCCATATCATCTCTGCTGATACGA	2147
QY	1021	CACATGGAATACCTCTGCTTGTCTACCAAGGTCACTTCGCTGGTTCAGTTCTTACTG	1080	QY	2101	GGGTGGACGCTGCTCTGAGGAGGAGCCGAAATTTCTCAAGTGTGTGATGAAGTTAGAA	2160
Db	1068	CACATGGAATACCTCTGCTTGTCTACCAAGGTCACTTCGCTGGTTCAGTTCTTACTG	1127	Db	2148	GGGTGGACGCTGCTCTGAGGAGGAGCCGAAATTTCTGAGGATTTCTCATGAGTTAGAA	2207
QY	1081	GATATGGAGCTGATATGATGATGCTGCTGATCCAGCAGCTCTAGTGGTGAAGAA	1140	QY	2161	GAGTGTCTCTGCAACATTTGAGCTGATCTCTCTGATCAAGTAAACAGCAGTGGGTCTCTC	2220
Db	1128	GATATGGAGCTGATATGATGATGCTGCTGATCCAGCAGCTCTAGTGGTGAAGAA	1187	Db	2208	GAGTGTCTCTGCAACATTTGAGCTGATCTCTCTGATCAAGTAAACAGCAGTGGGTCTCTC	2267
QY	1141	GATGACGACATGTTTGTATGGGCTTATGAAAGGGCATGTATGCCATTTGTCACACTC	1200	QY	2221	TCACCTCTCTTCTTCTGATGCTGTAACCGGGAGGAGCCTGGCGGAGTCACTGTG	2280
Db	1188	GATGACGACATGTTTGTATGGGCTTATGAAAGGGCATGTATGCCATTTGTCACACTC	1247	Db	2268	TCACCTCTCTTCTTCTGATGCTGTAACCGGGAGGAGCCTGGCGGAGTCACTGTG	2327
QY	1201	CTGAGACATTAAGAGACCAAGATGATGCTGCTGATGATGATTTCTCAGCCTCGGA	1260	QY	2281	GCAGCATTAAGAGAGTGGTTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGTGCT	2340
Db	1248	CTGAGACATTAAGAGACCAAGATGATGCTGCTGATGATGATTTCTCAGCCTCGGA	1307	Db	2328	GCAGCATTAAGAGAGTGGTTTTCGAATTTGGAATATGCTCTAAATGCAAGGTCCTATGTGCT	2387
QY	1261	GGAGATGGCTCTCTATGCTGTTCCATCACCTTTGGGGAAGATTTAAAGCATGACAAA	1320	QY	2341	TTGTCCAAAGTGTGACCAATATTTCTCTCAGAGTCTGCTTTGGAGGAGATGAAGA	2400
Db	1308	GGAGATGGCTCTCTATGCTGTTCCATCACCTTTGGGGAAGATTTAAAGCATGACAAA	1367	Db	2388	TTGTCCCAAGTGTGACCAATATTTCTCTCAGAGTCTGCTTTGGAGGAGATGAAGA	2447
				Db	2401	AGTCTTCAATACACACCACTTGAACAAATATGGCTATGTATCCGATCCCATGAGCTCAATG	2460

```
Db 2448 AGTCTTCATACACCCCATGACAAATATGGCTATGATCCCATGCCATGAGCTCAATG 2507
Qy 2461 CATTTTCATCTTCCGAAATAGTAGCAGCTTTGAGCAGCAGC 2505
Db 2508 CATTTTCATCTTCCGAAATAGTAGCAGCTTTGAGCAGCAGC 2552

RESULT 5
US-09-458-457-1
; Sequence 1, Application US/09458457
; Patent No. 6500654
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: mni-068cp
; CURRENT APPLICATION NUMBER: US/09/458,457
; CURRENT FILING DATE: 1999-12-10
; EARLIER APPLICATION NUMBER: 60/111,938
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/291,839
; EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2552)
US-09-458-457-1

Query Match 100.0%; Score 2505; DB 4; Length 3025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGAAATTAATACTAGACCAACCCAACTTGTAATGATGGAAGGAAAGGTC 60
Db 48 ATGGGAAATTAATACTAGACCAACCCAACTTGTAATGATGGAAGGAAAGGTC 107

Qy 61 AGTGAATCATATGTTATCAAAATAGAAAGATAGAGATGACCTGAGATCAAGGAAAA 120
Db 108 AGTGAATCATATGTTATCAAAATAGAAAGATAGAGATGACCTGAGATCAAGGAAAA 167

Qy 121 GAATGACAGAACTAAGGAATATTTGGCTCTGATGAAGGCTTCAGTAAAGTCAATT 180
Db 168 GAATGACAGAACTAAGGAATATTTGGCTCTGATGAAGGCTTCAGTAAAGTCAATT 227

Qy 181 AATTACCGCACTGAAATGGGCTGTCTTACTTCAATTTATGTTGCAATTTGCGAGGCA 240
Db 228 AATTACCGCACTGAAATGGGCTGTCTTACTTCAATTTATGTTGCAATTTGCGAGGCA 287

Qy 241 AATACATATTCGAATCTTATGTTGAAGGCTCCGCCATCTCGACTGACAGAAAT 300
Db 288 AATACATATTCGAATCTTATGTTGAAGGCTCCGCCATCTCGACTGACAGAAAT 347

Qy 301 GGAATTAACGCTTGCATTTAGCAGTTTACAGGATTAATGCAAGATTTGATCACTTCT 360
Db 348 GGAATTAACGCTTGCATTTAGCAGTTTACAGGATTAATGCAAGATTTGATCACTTCT 407

Qy 361 CTTACAGTGGAGTGNATATACAGCAGTTGATAGCGTGGCTGACCTCCATATTT 420
Db 408 CTTACAGTGGAGTGNATATACAGCAGTTGATAGCGTGGCTGACCTCCATATTT 467

Qy 421 GCTACAAATAGTGGCCACCTAGAGGCTGTGATGCTGTGCAACATGGAGCTAATGTC 480
Db 468 GCTACAAATAGTGGCCACCTAGAGGCTGTGATGCTGTGCAACATGGAGCTAATGTC 527

Qy 481 AATATCAAGATGAGTTTTTTTCACTCCATGCAATTTGCAAGGCTATGGAATGAA 540
Db 528 AATATCAAGATGAGTTTTTTTCACTCCATGCAATTTGCAAGGCTATGGAATGAA 587
```

```
Qy 541 CAGGTAACTCGCCTCTTTTGAATTTGGTGCTGATGTAAATGTAAAGTGGTGAAGTTGA 600
Db 588 CAGGTAACTCGCCTCTTTTGAATTTGGTGCTGATGTAAATGTAAAGTGGTGAAGTTGA 647
Qy 601 GATAGACCCCTCCACCTAGCATCTGAAAAGGATTTCTGAATATTTGCAAAATCTTTGATG 660
Db 648 GATAGACCCCTCCACCTAGCATCTGAAAAGGATTTCTGAATATTTGCAAAATCTTTGATG 707
Qy 661 GAAGAAGGAGCAAGCAGATGTGAATGCTCAAGATPAATGAAGACCATGCTCCACTCCAT 720
Db 708 GAAGAAGGAGCAAGCAGATGTGAATGCTCAAGATPAATGAAGACCATGCTCCACTCCAT 767
Qy 721 TTCTGTTCTCGATTTGGACACCATGATATAGTAAAGTATCTGCTGCAAAAGTGAATGGA 780
Db 768 TTCTGTTCTCGATTTGGACACCATGATATAGTAAAGTATCTGCTGCAAAAGTGAATGGA 827
Qy 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACCAT 840
Db 828 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTACCAT 887
Qy 841 GGCATAATTTGAAGTGGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
Db 888 GGCATAATTTGAAGTGGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947
Qy 901 GAAAAACATCTTCAGTGAACACAGCTTTTCATAGTGTGTACCTATGCGAAGACATTCAC 960
Db 948 GAAAAACATCTTCAGTGAACACAGCTTTTCATAGTGTGTACCTATGCGAAGACATTCAC 1007
Qy 961 CTAGTCAAAATTTCTTTGATCAGAAATGTCAATAACATCAACCAAGGAAAGGATGGG 1020
Db 1008 CTAGTCAAAATTTCTTTGATCAGAAATGTCAATAACATCAACCAAGGAAAGGATGGG 1067
Qy 1021 CACACTGGATTTACACTCTGCTGTACACAGCTGCANTTCGCTGGTTCAGTCTTACTTG 1080
Db 1068 CACACTGGATTTACACTCTGCTGTACACAGCTGCANTTCGCTGGTTCAGTCTTACTTG 1127
Qy 1081 GATAATGGAGCTCATATGAATCTTAGTGGCTTTGTATGCCAGCAGGTCTAGTGGTGA 1140
Db 1128 GATAATGGAGCTCATATGAATCTTAGTGGCTTTGTATGCCAGCAGGTCTAGTGGTGA 1187
Qy 1141 GATGAGCAGACATGTTGATGAGGCTTATGAAAAGGSCATGATGCCATTTGTCACTC 1200
Db 1188 GATGAGCAGACATGTTGATGAGGCTTATGAAAAGGSCATGATGCCATTTGTCACTC 1247
Qy 1201 CTGAAGCATTTAAGAGACACACAGATGAATTCGCTGTAATGAATATTTCTCAGCTCGA 1260
Db 1248 CTGAAGCATTTAAGAGACACACAGATGAATTCGCTGTAATGAATATTTCTCAGCTCGA 1307
Qy 1261 GAGATGGCTCTCATGTTGCTTCCATCACCTTTGGGAAAGATTAAGACATGACAAA 1320
Db 1308 GAGATGGCTCTCATGTTGCTTCCATCACCTTTGGGAAAGATTAAGACATGACAAA 1367
Qy 1321 GAGAAGCAGATATTTCTCTTAAGAGCTGGATTCCTTACATTTCCATCTTCAGCTC 1380
Db 1368 GAGAAGCAGATATTTCTCTTAAGAGCTGGATTCCTTACATTTCCATCTTCAGCTC 1427
Qy 1381 TCAGAAATTTGAGTTCCATGAGATTTAGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
Db 1428 TCAGAAATTTGAGTTCCATGAGATTTAGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1487
Qy 1441 CGATGAGAAATAAATAGTGGCTATAAACCCTTATCGAGCCATACCTACTGCTCCAG 1500
Db 1488 CGATGAGAAATAAATAGTGGCTATAAACCCTTATCGAGCCATACCTACTGCTCCAG 1547
Qy 1501 TCAGATGTGGATATGTTTGGCGAGAGGTGTCATTTCTCTGACAGTCAATCATCCCTGC 1560
Db 1548 TCAGATGTGGATATGTTTGGCGAGAGGTGTCATTTCTCTGACAGTCAATCATCCCTGC 1607
Qy 1561 GTAATTCAGTTTGTGGTCTCTTGTGAATGATCCAGCCAGTTTGGCATTTGTCTACTCA 1620
Db 1608 GTAATTCAGTTTGTGGTCTCTTGTGAATGATCCAGCCAGTTTGGCATTTGTCTACTCA 1667
```


Db 828 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGCATGCTACAA 887
Qy 841 GGCAGAAATTTGAAGTSCCAGGAATCATCAATATCAGGACAGAAAGTCTGACTAAG 900
Db 888 GGCAGAAATTTGAAGTSCCAGGAATCATCAATATCAGGACAGAAAGTCTGACTAAG 947
Qy 901 GAAACATCTTCAGTGAACAGCTTTTCATAGTGTGTACCTATGCGCAAGAGCATTCAC 960
Db 948 GAAACATCTTCAGTGAACAGCTTTTCATAGTGTGTACCTATGCGCAAGAGCATTCAC 1007
Qy 961 CTAGTCAAAATTTCTTTGATCAGATGTCAATAAACATCAACCAAGAGGAGGATGGG 1020
Db 1008 CTAGTCAAAATTTCTTTGATCAGATGTCAATAAACATCAACCAAGAGGAGGATGGG 1067
Qy 1021 CACACTGATTTACACTCTGCTGCTACCAAGGTTCATTCGCTGCTTCAGTTCCTTACTG 1080
Db 1068 CACACTGATTTACACTCTGCTGCTACCAAGGTTCATTCGCTGCTTCAGTTCCTTACTG 1127
Qy 1081 GATAATGGAGCTGATATGAATAGTGGTGTGATCCAGCAGGTCATAGTGGTGAAGAA 1140
Db 1128 GATAATGGAGCTGATATGAATAGTGGTGTGATCCAGCAGGTCATAGTGGTGAAGAA 1187
Qy 1141 GATGAGCAGACATGTTGATGTGGCTTATGAAAGGGCATGATCCCATTCACACTC 1200
Db 1188 GATGAGCAGACATGTTGATGTGGCTTATGAAAGGGCATGATCCCATTCACACTC 1247
Qy 1201 CTGAAGCATTTAAGAGACACCAAGATGAATTCCTGTAATGAATATTCCTCAGCTCGA 1260
Db 1248 CTGAAGCATTTAAGAGACACCAAGATGAATTCCTGTAATGAATATTCCTCAGCTCGA 1307
Qy 1261 GGAGATGGCTCTATGCTGTCTGCTCCATCCCTTGGGAGAGATTAAGAGCATGACAAA 1320
Db 1308 GGAGATGGCTCTATGCTGTCTGCTCCATCCCTTGGGAGAGATTAAGAGCATGACAAA 1367
Qy 1321 GAGAGGCAGATATTCCTCCTTAAGAGCTGATGCTTCATATTCATCTCAGCTC 1380
Db 1368 GAGAGGCAGATATTCCTCCTTAAGAGCTGATGCTTCATATTCATCTCAGCTC 1427
Qy 1381 TCAGAAATGAGTTCATGATATATGCTCAGTTCCTTTGGGAAAGTATATAAGGA 1440
Db 1428 TCAGAAATGAGTTCATGATATATGCTCAGTTCCTTTGGGAAAGTATATAAGGA 1487
Qy 1441 CGATGCGAATATAATAGTGGCTTATAAGAGTATCGAGCAATACCTACTGCTCCAG 1500
Db 1488 CGATGCGAATATAATAGTGGCTTATAAGAGTATCGAGCAATACCTACTGCTCCAG 1547
Qy 1501 TCAGATGTGATATGTTTTCGCGAGAGTGTCCATTCCTGCGAGCTCAATCATCCCTGC 1560
Db 1548 TCAGATGTGATATGTTTTCGCGAGAGTGTCCATTCCTGCGAGCTCAATCATCCCTGC 1607
Qy 1561 GTAAATCAGTTTGGGTGCTTGTGATGATCCAGCCAGTTGCCATTTGCTACTCAA 1620
Db 1608 GTAAATCAGTTTGGGTGCTTGTGATGATCCAGCCAGTTGCCATTTGCTACTCAA 1667
Qy 1621 TACATATCAGGGGTCTCTGTTCTCCCTCTTCATGAGCAGAGAGATTTCTGATTTG 1680
Db 1668 TACATATCAGGGGTCTCTGTTCTCCCTCTTCATGAGCAGAGAGATTTCTGATTTG 1727
Qy 1681 CAGTCTAAATTAATATTGAGTAGATGTTGCCAAAGGATGAGTACCTTCACAACTGG 1740
Db 1728 CAGTCTAAATTAATATTGAGTAGATGTTGCCAAAGGATGAGTACCTTCACAACTGG 1787
Qy 1741 ACAGAGCAATATACATCTGAGTGAACAGTCAATATTCCTCTATGAGGATGGG 1800
Db 1788 ACAGAGCAATATACATCTGAGTGAACAGTCAATATTCCTCTATGAGGATGGG 1847
Qy 1801 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTCAGTCTCTGGATGAAGCAAC 1860
Db 1848 CATGCTGTGTGGCAGATTTTGGAGATCAAGATTTCTCAGTCTCTGGATGAAGCAAC 1907
Qy 1861 ATGACAAAACAACTGGGACCTCCGTTGGATGGCTCCTGAGGTGTTACGAGTGCCT 1920

Db 1908 ATGACAAAACAACTGGGAACTCCGTTGGATGGCTCTCGAGGTGTTCCGCGAGTGCCT 1967
Qy 1921 CGGTACACCATCAAGCAGATGTTTTCAGCTATGCTCTGTGCTCTGTGGAAATTTCTCACT 1980
Db 1968 CGGTACACCATCAAGCAGATGTTTTCAGCTATGCTCTGTGCTCTGTGGAAATTTCTCACT 2027
Qy 1981 GGGGAAATTTCCATTCGCTCATCTCAAGCCAGCGGCTGCGGACAGACATGCTTACCAC 2040
Db 2028 GGGGAAATTTCCATTCGCTCATCTCAAGCCAGCGGCTGCGGACAGACATGCTTACCAC 2087
Qy 2041 CACATCAGACCTCCCATTTGGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2100
Db 2088 CACATCAGACCTCCCATTTGGCTATTCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2147
Qy 2101 GGGTGGAAACGATGCTCTGAAAGGAAGACCCGAAATTTTCTGAAAGTTGTCATGAAGTTAGAA 2160
Db 2148 GGGTGGAAACGATGCTCTGAAAGGAAGACCCGAAATTTTCTGAAAGTTGTCATGAAGTTAGAA 2207
Qy 2161 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2220
Db 2208 GAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGATCAAGTAAACAGAGTGGTCTCTC 2267
Qy 2221 TCACCTTTCTTTCTTCTGATTTGCCCTGTTGAACCGGGGAGGACCTGGCGGAGTCAATGG 2280
Db 2268 TCACCTTTCTTTCTTCTGATTTGCCCTGTTGAACCGGGGAGGACCTGGCGGAGTCAATGG 2327
Qy 2281 GCGCATTAAGAAAGTCTGTTTGAATTTGAATTTGCTCTAAATGCAAGGTCCTATGCTGCT 2340
Db 2328 GCGCATTAAGAAAGTCTGTTTGAATTTGAATTTGCTCTAAATGCAAGGTCCTATGCTGCT 2387
Qy 2341 TTGTCCCAAGTCTCGCAATATTTCTCTCAAGTCTCTCTTTGGAGGAGATGAAAGA 2400
Db 2388 TTGTCCCAAGTCTCGCAATATTTCTCTCAAGTCTCTCTTTGGAGGAGATGAAAGA 2447
Qy 2401 AGTCTTCAATACACACCCCATTTGCAATATTTGCTATGCTATCCGATCCCATGAGCTCAATG 2460
Db 2448 AGTCTTCAATACACACCCCATTTGCAATATTTGCTATGCTATCCGATCCCATGAGCTCAATG 2507
Qy 2461 CATTTTCATTTGCGGAAATAGTAGCAGCTTTTGAGGACAGCAGC 2505
Db 2508 CATTTTCATTTGCGGAAATAGTAGCAGCTTTTGAGGACAGCAGC 2552

RESULT 7

US-09-458-457-9
; Sequence 9, Application US/09458457
; Patent No. 6500654
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: mni-068cp
; CURRENT APPLICATION NUMBER: US/09/458,457
; CURRENT FILING DATE: 1999-12-10
; EARLIER APPLICATION NUMBER: 60/111,938
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/291,839
; EARLIER FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-458-457-9

Query Match 75.5%; Score 1892.2; DB 4; Length 2505;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;


```
Db 2161 GAGTGCCTATCAATGTGGAGCTCATGTCTCCAGCATCAAGTAACAGCAGTGGCTCTCTG 2220
Qy 2221 TCACCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 2280
Db 2221 TCACCTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 2280
Qy 2281 GCAGCATTAGAAAGTGGTTCGAAATTTGGAATATGCTCTAAATGCAAGGTCTCTATGCTGCT 2340
Db 2281 GCAGCCTTACGGAGCGTTTGGATTGGAGTATGCCCTAAATGCAAGGTCTCTATGCTGGG 2340
Qy 2341 TTGTCCCAAGTCTGGACAATATTCCTCTCAAGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 2400
Db 2341 TGGTCCCAAGTCTGGACAATATTCCTCTCAAGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 2400
Qy 2401 AGTCTTCAATACACACCATTTGCAATATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 2460
Db 2401 AGCACCAGTATTCAACTGTGCAATATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 2460
Qy 2461 CATTTTCATTTTCCGCAATATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 2505
Db 2461 CACCTTCACTCCCGCCCAAGACGACGCAACTTTGAGACAGCAAC 2505

RESULT 8
US-09-947-199A-9
; Sequence 9, Application US/09947199A
; Patent No. 6660490
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-068CP2
; CURRENT FILING DATE: 2001-09-05
; PRIOR FILING DATE: US 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199A-9

Query Match 75.5%; Score 1892.2; DB 4; Length 2505;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 1 ATGGGAAATTAATAATCTAGCAACCCAACTTGTCTGATGATGGAAGAAAAGTC 60
Db 1 ATGGGAAATTAATAATCTAGCAACCCAACTTGTCTGATGATGGAAGAAAAGTC 60
Qy 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTGAGATCAAGAAAAA 120
Db 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAGATGACCTGAGATCAAGAAAAA 120
Qy 121 GAATGACAGACTAGGAATATATTTGCTGCTGATGAAGCCCTTCAGTAAAGTCAATTA 180
Db 121 GAATTTCAAGAACTAGGAATATATTTGCTGCTGATGAAGCCCTTCAGTAAAGTCAATTA 180
Qy 181 AATTACCGCACTAGGAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 181 AATTACCGCACTAGGAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Qy 241 AATATCATATTCGAATCTTTATGTTGAAAGGGCTCCGCGCAATCTCGACTGACAGAAAT 300
Db 241 AATATCATATTCGAATCTTTATGTTGAAAGGGCTCCGCGCAATCTCGACTGACAGAAAT 300
```

```
Db 241 AAGTCACATATCCGTGCCCTTATGTTAAAGGGCTCCGTCATCCAGACTGACGAGAAAT 300
Qy 301 GGATTTACAGCCTTGCAATTTACAGATTTTAAAGGATTAATGCAAGATTTGATCATTCTCTG 360
Db 301 GGGTTTTCCAGCTCTGCACTCGCCCGTTTACAAAGGACAGCCCGGAACTTATCATTCTACTG 360
Qy 361 CTTTCAGATGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCTCACTGCCCTCCATATT 420
Db 361 TTGCAAGCGGAGCAGATGTTTACGACAGTGGATACGGTGGCTCTCAAGCCCTCCACATA 420
Qy 421 GCTTCAATAGCTGGCCACCTAGAGCTGCTGATGCTGTTGTCACATGAGAGCTTAATGTC 480
Db 421 GCTGCAATAGCTGGACACCCAGAGGCTGCAAGATGCTGCTACACACATGAGGCAATGTC 480
Qy 481 AATATTCAAGATGCAAGTCTTTTTCATCCATTCATTCATTCATTCATTCATTCATTCATTC 540
Db 481 AATGTTCAAGATGCGCTCTTCTTACCACCTGCAATTCAGCCCTACTATGCGCACGAG 540
Qy 541 CAGGTAACCTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA 600
Db 541 CAGGTAACCTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA 600
Qy 601 GATAGACCCCTCCACTAGCATCTGCAAAAAGGATTTTGAATATTGCAAAACTCTTGATG 660
Db 601 GACAGCCCTCTGCACCTGGCCTCTGCAAGGGCTTCTTCAACATTTGTGAAACTCTTGTA 660
Qy 661 GAAGAAGCAGCAAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTGCCACTCCAT 720
Db 661 GAAGAAGCAGCAAAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTGCCACTCCAT 720
Qy 721 TTTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATTTGAA 780
Db 721 TTTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATTTGAA 780
Qy 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACCTTACCTTGGCATGTACAAT 840
Db 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACCTTGGCATGTACAAT 840
Qy 841 GGCATAATTTGAAGTTGCCAAGGAATATCCAAATPATCAGGAACAGAAAGTCTGACTAAG 900
Db 841 GGCATAATTTGAAGTTGCCAAGGAATATCCAAATPATCAGGAACAGAAAGTCTGACTAAG 900
Qy 901 GAAACATCTTCAGTGAACAGCTTTTCATAGTGTCTGCTACCTATGCTGCAAGAGATGAC 960
Db 901 GAAACATCTTCAGTGAACAGCTTTTCATAGTGTCTGCTACCTATGCTGCAAGAGATGAC 960
Qy 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCATAAATCATCAACCAAGGAAGGATGGG 1020
Db 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCATAAATCATCAACCAAGGAAGGATGGG 1020
Qy 1021 CACACTGGATTAACATCTGCTTGTACCAAGTCAATTCGCTGGTTCAGTCTTACTG 1080
Db 1021 CACACTGGATTAACATCTGCTTGTACCAAGTCAATTCGCTGGTTCAGTCTTACTG 1080
Qy 1081 GATAATGGAGCTGATATGAATCTAGTGTCTGTATCCAGCAGGTCTAGTGGTGAATAA 1140
Db 1081 GATAATGGAGCTGATATGAATCTAGTGTCTGTATCCAGCAGGTCTAGTGGTGAATAA 1140
Qy 1141 GATGACAGACATGTTGATGCTGCTTATGAAAAGGGCATGATGCCATTGTCACTC 1200
Db 1141 GATGACAGACATGTTGATGCTGCTTATGAAAAGGGCATGATGCCATTGTCACTC 1200
Qy 1201 CTGAGCATTTAAGAGACCAAGATGAATTCCTGCTGATGATGATTTCTCAGCCTGGA 1260
Db 1201 CTGAGCATTTAAGAGACCAAGATGAATTCCTGCTGATGATGATTTCTCAGCCTGGA 1260
Qy 1261 GGAGATGGCTCTATGCTGCTTCCATCACCTTTGGGAAAGATTTAAAGCATGACAAA 1320
Db 1261 GGAGATGGCTCTATGCTGCTTCCATCACCTTTGGGAAAGATTTAAAGCATGACAAA 1320
Qy 1321 GAGAAGGAGATATTTCTCTCTCTAAGAGCTGATGCTTCAATTTCCATCTTCAGCTC 1380
Db 1321 GAGAAGGAGATATTTCTCTCTCTAAGAGCTGATGCTTCAATTTCCATCTTCAGCTC 1380
```

```
QY 1381 TCAGAAATTGAGTTCATGATTAATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
Db 1381 TCCGAAATCGAGTTCACGAGATTATCGGCTCGGTTCTTTGGGAAAGTCTATARAAGG 1440
QY 1441 CGATGAGAAATAAATAGTGGCTATAAAAGTTATCGAGCCAAATACCTACTGCTCCAAG 1500
Db 1441 CGATGAGAAATAAATAGTGGCTATAAAAGTTATCGAGCCAAATACCTACTGCTCCAAG 1500
QY 1501 TCAGATGAGATATGTTTGGCCGAGAGGTGTCCATCTCTGCCAGTCAATCATCCCTGC 1560
Db 1501 TCAGACGTGGATATGTTTGGCCGAGAGGTGTCCATCTCTGCCAGTCAATCATCCCTGC 1560
QY 1561 GTAATTCAGTTTGGGTGCTTGTGAATGATCCAGCCAGTTTGCCTACTCAAA 1620
Db 1561 GTGGTTCAAGTTTGGGTGCTTGTGAATGATCCAGCCAGTTTGCCTACTCAAG 1620
QY 1621 TACATATCAGGGGTTCTGTTCTCCCTCTTCATGAGCAGAGAGGATCTTGATTG 1680
Db 1621 TACATTTCCAGAGGCTCCCTGTTCTCCCTGCTTCATGAACAGAGAGAAATCTTGACTTG 1680
QY 1681 CAGTCTAAATTAATTCAGTATGATGTTGCCAAAGGCATGGAGTACCTTCACAACTTG 1740
Db 1681 CAGTCTAAATTAATTCAGTATGATGTTGCCAAAGGCATGGAGTACCTTCACAGCTTG 1740
QY 1741 ACACAGCAATATACATCGTGTGATGAACAGTCAATATTCCTCTATGAGGATGGG 1800
Db 1741 ACCAGCAATATACATCGTGTGATGAACAGTCAATATTCCTCTATGAGGATGGG 1800
QY 1801 CATGCTGTGGTGGAGATTTGGAGATCAAGATTCTACAGTCTCTGGATGAAGACAAC 1860
Db 1801 CATGCTGTGGTGGAGATTTGGAGATCAAGATTCTACAGTCTCTGGATGAAGACAAC 1860
QY 1861 ATGACAAACCACTGGGAACCTCCGTTGGATGGCTCCTCAGTGTTCACGAGTGCAT 1920
Db 1861 ATGACAAAGCAGCCAGGAACTCGCTGGATGGCCCTGAGTGTTCACAGTGCAGG 1920
QY 1921 CGGTACACATCAAGCAGATGCTTCAGCTATGCTGTCTGTCTGGGAAATTCATCT 1980
Db 1921 AGATACACATCAAGCAGTGTCTTCAGTTCCTCTGTCTGTGGAGTCTCTCAT 1980
QY 1981 GCGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGGCGAGCAGACATGGCTTACCAC 2040
Db 1981 CGAGAAATTCATTCGCTCATCTCAAGCCAGCGGCTCGGCGAGCAGATGGCGTATCAC 2040
QY 2041 CACATCAGACCTCCATGGCTATTCATTCGAGGAGCAGATCATCTCTGCTGATACGA 2100
Db 2041 CACATCAGACCGCCATCGGCTATTCATTCGAGGAGCAGATCATCTCTGCTGATACGG 2100
QY 2101 GGGTGAACGATGCTCTGAAGGAGACCCGAAATTTCTGAAGTTGTATGAAGTTAGAA 2160
Db 2101 GGTGAGATGATGCTCTGAAGGAGACCCGAGATTTCTGAGTGTGTAGCAAACTGGAG 2160
QY 2161 GAGTGTCTCTGCAATGAGTGTCTCTGATCAAGTAAAGTAAAGCAGTGGGTCTCTC 2220
Db 2161 GAGTGTCTCTGCAATGAGTGTCTCTGATCAAGTAAAGTAAAGCAGTGGGTCTCTC 2220
QY 2221 TCACCTTCTCTCTCTCTGATGCTGTAACCGGGAGGAGCTGCGCGAGTCAATGTG 2280
Db 2221 TCACCTTCTCTCTCTCTGATGCTGTAACCGGGAGGAGCTGCGCGAGTCAATGTG 2280
QY 2281 GCAGATTAAGAAGTCGTTTGAATATGCTCTAAATGCAAGGTCCTATGCTGT 2340
Db 2281 GCAGCTTACGAGCGGTTTGAATATGCTCTAAATGCAAGGTCCTATGCTGT 2340
QY 2341 TTGTCGAAAGTCTGGAATATTCCTCAAGGTCGTCTTTGGAGGAGTGAAGA 2400
Db 2341 TGGTCCCAAGGTTTGGAAACACACTCTAAATCCGGGCTGTCTTTGGAGGAGTGAATAG 2400
QY 2401 AGTCTTCAATACACACCCATATGATGGCTATGATCCGATCCCATGAGTCAATG 2460
Db 2401 AGCACCAGTATTCAACTGTTGACAAATACGGCTATGTGCTGATCCCATGAGCTGACG 2460

Query Match 75.5%; Score 1892.2; DB 4; Length 3026;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

QY 1 ATGGGAAATTAATACTAGACCAACCCAAACTTGTATGATGAATGGAAGAAAGATC 60
Db 61 ATGGGAAATTAATACTAGACCAACCCAAACTTGTATGATGAATGGAAGAAAGATC 120
QY 61 AGTGAATCATATGTTATCAATAGAAAGATAGAGATGACCTCAGATCAAGCAAAA 120
Db 121 AGTGAATCATATGTTATCAATAGAAAGATAGAGATGACCTCAGATCAAGCAAAA 180
QY 121 GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180
Db 181 GAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 240
QY 181 AATTACCGGACTGAAATGGGCTGTCTCTACTTCACTTATGTTGATTTGTTGGAGGCAAG 240
Db 241 AATTACCGGACTGAAATGGGCTGTCTCTACTTCACTTATGTTGATTTGTTGGAGGCAAG 300
QY 241 AATTCACATATTCGAACTCTTATGTTGAAAGGCTCCGCCATCTCGACTGACAAAGAAAT 300
Db 301 AAGTCAATATTCGTCGCCCTTATGTTTAAAGGCTCCGTCATCCAGACTGACGAGAAAT 360
QY 301 GGATTTACGCTTGAATTTAGCAGTTTACAGGATATGAGAAATGATCACTTCTCTG 360
Db 361 GGGTTCCAGCTCTGCACTGGCCGTTTACAGGACAGCCGGAACCTTATCACTTCACTG 420
QY 361 CTTTACAGTGGAGTATATACAGAGGTTGGATAGCGTGGCTCACTGCCCTCCATATT 420
Db 421 TTGCACAGCGGAGAGATGTTTACAGAGTGGATAGCGTGGCTCACAGCTCCACATA 480
QY 421 GCTCAATAGCTGCCCACTTAGAGGCTGTGATGTTGCTGTTGCAACATGGAGCTAATGTC 480
Db 481 GCTCAATAGCTGCCCACTTAGAGGCTGTGATGTTGCTGTTGCAACATGGGCGCAATGTG 540
QY 481 AATTTCAAGATCAGTGTGTTTTCCTCCATTCATTCAGTGAATGATGATGATGATGAA 540
Db 541 AATTTCAAGATCAGTGTGTTTTCCTCCATTCATTCAGTGAATGATGATGATGATGAA 600
QY 541 CAGTAACTCGCCTTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTCAAGTGGTA 600
```

Db 601 CAGGTAAACAGTGTCTCTTTTGAAGTTTGGTGTGATGCTCAATGTAAAGCGGTGAAGTTGGG 660
Qy 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTCTTGAATTAATTCAAAACTCTTGATG 660
Db 661 GACAGGCTCTGCACTGCGCTCTGCAAAAGGCTCTTCAACANTGTGAAACTCTGCTGTA 720
Qy 661 GAGAGGAGCAAGACAGATGTGATGCTCAAGATATGAGACCATGTCCCTCCAT 720
Db 721 GAAGAAGGAGCAAGACAGATGTGATGCTCAAGACCATGTGAGACCATGTCCCTCTGAC 780
Qy 721 TTCTGTTCTCGATTGTCGACCAACATGATAGTTAAAGTATCTCTGCAAAAGTGAATTCGAA 780
Db 781 TTCTGTTCTCGATTGTCGACCAACATGATAGTTAAAGTATCTCTGCAAAAGTGAATTCGAA 840
Qy 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGACCTGGCATCTACAT 840
Db 841 GTCCAGCTCAGCTCATTAACATCTATGTTGACACTCTCTTTCAGCTGGCATCTACAT 900
Qy 841 GGCATAATTCAGTTGCGCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
Db 901 GGAATTTTGAAGTTGCCAAGGAAATGTCAGGTAAACAGGAACTGAAAGTCTGACTAAG 960
Qy 901 GAAAAATCTTCAGTGAACAGCTTTTCAATAGTGTGATACCTATGCGCAAGCAATTCAC 960
Db 961 GAAAAATCTTCAGTGAACAGCTTTTCAATAGTGTGATACCTATGCGCAAGCAATTCAC 1020
Qy 961 CTAGTCAAAATTTCTTGTGATGAGATGCTAAAGATCAACCAACCAAGGAGGATGGG 1020
Db 1021 CTGGTCAAAATTTCTTGTGATGAGATGCTGAAATTAACCAACCGAAGAGATGGG 1080
Qy 1021 CACACTGGATTACACTCTGCTTCTACCAAGCTTCAATTCGCTGGTTTCAAGTTCTTACTG 1080
Db 1081 CACACAGGATTGCACTCTGCTTCTACCAAGCTTCAATTCGCTGGTTTCAAGTTCTTACTG 1140
Qy 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTATGCCAGAGCTCTAGTGGTGAAGAA 1140
Db 1141 GATAATGGTGCAGATATGAATCTTGTGCTTGTGATCCAGAGCTCTAGTGGTGAAGAA 1200
Qy 1141 GATGAGCAGACATGTTTGTATGGGCTTATGAAAGGCGCATGATGCCATGTCACACTC 1200
Db 1201 GATGAGCAGACATGTTTGTATGGGCTTACGAGAAGGACATGATGCCATGTTTACACTC 1260
Qy 1201 CTGAAGCATTTAAGAGACCAAGATGAATGTCCTGTAATGAATTTCTCAGCTCGA 1260
Db 1261 CTGAAGCATTTAAGAGACCAAGATGAATGTCCTGTAATGAATTTCTCAGCTCGA 1320
Qy 1261 GGAGATGGCTCCTATGCTGTGCTTCCATCACCTTGGGAGATTAAGAGCATGACAAA 1320
Db 1321 GGAGATGGCTCCTATGCTGTGCTTCCCTTGGGCAAGATTAAGAGCATGACAAA 1380
Qy 1321 GAGAAGCAGATATTTCTCTCTTAAGAGCTGGATGCTTCAATTTCCATCTTCAGCTC 1380
Db 1381 GAGAAGCAGATATTTCTCTCTTAAGAGCTGGATGCTTCAATTTCCATCTTCAGCTC 1440
Qy 1381 TCAGAAATGAGTTCCATGAGATATTTGGTCAAGTTCTTTGGGAAAGTATTAAGAA 1440
Db 1441 TCCGAAATCGAGTTCCACGAGATTAATCGGCTCGGGTTCTTTTGGGAAAGTATTAAGAGG 1500
Qy 1441 CGATGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTTACTGCTCCAAAG 1500
Db 1501 CGATGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTTACTGCTCCAAAG 1560
Qy 1501 TCAGATGTGATATGTTTCCGAGAGGTGTCATTTCTCTGTCAGCTCAATCATCTCCCTGC 1560
Db 1561 TCAGATGTGATATGTTTCCGAGAGGTGTCATTTCTCTGTCAGCTCAATCATCTCCCTGC 1620
Qy 1561 GTAATTCAGTTTGGGTGCTTCTGTAATGATCCAGCAGTTTCCCATTTGCTACTCAA 1620
Db 1621 GTGTTTCAGTTTGGGTGCTTCTGTAATGATCCAGCAGTTTCCCATTTGCTACTCAA 1680
Qy 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTTCAATGAGCAGAGAGGATTTCTGATTTG 1680
Db 1681 TACATTTTCAGAGGCTCCCTGTTCTCTCTCTCTCATGACAGAGAGATTTCTGACTTG 1740

Qy 1681 CAGTCTAAATTAATTAATTTGCAGTAGATGTTGCCAAGGATGGAGTACCTTCACAACTG 1740
Db 1741 CAGTCTAAATTAATTAATTTGCAGTAGATGTTGCCAAGGATGGAGTACCTTCACAACTG 1800
Qy 1741 ACACAGCCAAATTAATATCATGTCAGTTGAACAGTCACAATATTTCTTCTCTATGAGGATGG 1800
Db 1801 ACCCAGCCAAATTAATATCATGTCAGTTGAACAGTCACAATATTTCTTCTCTATGAGGATGG 1860
Qy 1801 CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTTACAGATCTCTGGATGAAGACAAC 1860
Db 1861 CATGCTGTGGTGGCAGATTTTGGAGATCAAGATTTTACAGATCTCTGGATGAAGACAAC 1920
Qy 1861 ATGACAAAAACACCTGGGAACTCTGTTGGATGGCTCTGAGGTGTTTACGCAAGTCACT 1920
Db 1921 ATGACAAAAACACCTGGGAACTCTGTTGGATGGCTCTGAGGTGTTTACGCAAGTCACT 1980
Qy 1921 CGGTACACCATCAAAAGCAGATGTTTCAAGTATGCTGTTGTGTTGGGAAATTTCTACT 1980
Db 1981 AGATACACCATCAAAAGCAGATGTTTCAAGTATGCTGTTGTGTTGGGAAATTTCTACT 2040
Qy 1981 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGCTGCGGAGCAGACATGGCTTACCAC 2040
Db 2041 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGCTGCGGAGCAGACATGGCTTACCAC 2100
Qy 2041 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCAAAGCCCATATCATCTCTCTGATACGA 2100
Db 2101 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCAAAGCCCATATCATCTCTCTGATACGG 2160
Qy 2101 GGTGGAACGATGCTCTCAAGGAGAACCCGAAATTTTCTGAAGTTGTCATGAAGTTAGAA 2160
Db 2161 GGTGGAACGATGCTCTCAAGGAGAACCCGAAATTTTCTGAAGTTGTCATGAAGTTAGAA 2220
Qy 2161 GAGTGTCTCTGCAACATTCAGCTGATGCTCTCATCAAGTAAACAGCAGTGGGTCTCTC 2220
Db 2221 GAGTGTCTCTGCAACATTCAGCTGATGCTCTCATCAAGTAAACAGCAGTGGGTCTCTC 2280
Qy 2221 TCACCTTCTCTTCTTCTGATTTGCTGTAACCCGGGAGGACCTGGCCGAGGATCATGTG 2280
Db 2281 TCACCTTCTCTTCTTCTGATTTGCTGTAACCCGGGAGGACCTGGCCGAGGATCATGTG 2340
Qy 2281 GCAGCATTAAGAGTGGTTTTCGAAATGGAATGCTCTAAATGCAAGTCTCTATCTCTGCT 2340
Db 2341 GCAGCATTAAGAGTGGTTTTCGAAATGGAATGCTCTAAATGCAAGTCTCTATCTCTGCT 2400
Qy 2341 TTCTCCAAAGTCTCGCAATTAATTTCTCTCAAGTCTGTCTTTGGAGGATGAAAGA 2400
Db 2401 TGCTCCAAAGTCTCGCAATTAATTTCTCTCAAGTCTGTCTTTGGAGGATGAAAGA 2460
Qy 2401 AGTCTTCAATACACCCCATTTGACAAATATGCTATGATCCGATCCCATGAGCTCAATG 2460
Db 2461 AGCACCAGTATTCAACTGTTGACAAATACGGCTATGTTCTGTATCCCATGAGCTGAGC 2520
Qy 2461 CATTTTCAATTTTGGCGAAATAGTAGCAGCTTTTGAGGACAGCAGC 2505
Db 2521 CACCTTCACTCCCGCCCAAGCAGCAGCACTTTTGAGGACAGCAAC 2565

RESULT 10

US-09-947-199A-7

; Sequence 7, Application US/09947199A

; Patent No. 6660490

; GENERAL INFORMATION:

; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: MNI-068CP2

; CURRENT APPLICATION NUMBER: US/09/947,199A

; CURRENT FILING DATE: 2001-09-05

; PRIOR APPLICATION NUMBER: US 60/111,938

; PRIOR FILING DATE: 1998-12-11

; PRIOR APPLICATION NUMBER: US 09/291,839

; PRIOR FILING DATE: 1999-04-14

1981	AGATACACCATCAAGGCTGATGTCCTTCAGTTACTCCCTGTGTCTGTGGAGACTCCTCACT	2040
Qy		
1981	GGCGAAATTCATTTCGCTCATCTCAAGCCAGCGGCTGGCGGACGACAGATGGCTTACCAC	2040
Db		
2041	GGAGAAATTCATTTCGCTCATCTCAGCCAGCGCTGCAGACGACAGATATGGCGTATCAC	2100
Qy		
2041	CACATCAGACCTCCATTCGCTATTTCATTCGACGCCATATCATCTCTCTGATACGA	2100
Db		
2101	CACATCAGACGCCATCGGCTATTTCATCCCCAAGCCCATCTCATCCCTCTGATACGG	2160
Qy		
2101	GGGTGGACGACATCTCTGAAGGAGACCGGAAATTTCTGAAGTTGTCTATCAAGTTAGAA	2160
Db		
2161	GGCTGGAAATGCATGTCTGAAGACGACAGAGTTCTCTGAAGTCTTTAGCAAACTGGAG	2220
Qy		
2161	GAGTGTCTCTGCACAAATTCGAGCTGATGTCTCTTGATCAAGTAAACAGCAGTGGGTCTCTC	2220
Db		
2221	GAGTGCCTATGCAATTCGAGGCTCATGTCTCCAGCATCAAGTAAACAGCAGTGGGTCTCTG	2280
Qy		
2221	TCACCTCTCTTCTTCTGATTCGCTGGTGAACCGGGGAGGACCTGGCGGAGTCATGTG	2280
Db		
2281	TCACCTCTCTTCTTCCGATTTCGCTGAGCGGGGAGGCGCTGGCGGAGCCACGTG	2340
Qy		
2281	GCAGCATTAAGAAATCGTTTCGAATTCGAATATGCTCTAAATGCAAGTCTCTATGCTGCT	2340
Db		
2341	GCAGCTTACGGAGCCGTTTGAGTTGGAGTATGCCCTAAATGCAAGTCTCTATGCTGGG	2400
Qy		
2341	TTGTCCCAAGTGTGGACAATATTCCTCTCAAGGTCTCTTTTGGAGGAGATGAAAAGA	2400
Db		
2401	TGGTCCCAAGTGTGGAAACACACTCTAAATCCGGGCGCTCTTTTGGAGGAGATGAATAGG	2460
Qy		
2401	AGTCTTCAATACACACCATTTGACAAATATGGCTATGTATCCGATCCCATGAGTCAATG	2460
Db		
2461	AGCACCCAGTATTCAAATGTTGCAAAATACGCGCTATGTCTGTATCCCATGAGCGCTGACG	2520
Qy		
2461	CATTTTTCATTCTCCGAAATAGTAGCAGCTTTGAGGACAGCAGC	2505
Db		
2521	CACCTTCACTCCCGCAAGACGACAGCAACTTTTGGAGCAGCAAC	2565

```

RESULT 11
US-09-833-381-1149
; Sequence 1149, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1149
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1149

```

	Query Match	17.1%;	Score 427.4;	DB 4;	Length 616;
	Best Local Similarity	81.8%;	Pred. No. 3.1e-124;		
	Matches 505;	Conservative 0;	Mismatches 111;	Indels 1;	Gaps 1;
QY	291	GACAGAAATGGATTACAGCCTTGCAATTTACAGTTTACAGGATAATGCAGAAATGAT	350		
DB	1	GACGAGAAATGGTTTCCAGTCTGSCACCTGGCCGTTTACAGGACAGCCCGGAACCTTAT	60		
QY	351	CACCTTCTCTGCTTACAGTGGAGCTGTATATACAGCAGGTTGGATACGGTGGGCTC	410		
DB	61	CACCTTCAGTTGTCACACGGAGCAGATGTTTCAGCAAGTGGGATACGGTGGGCTC	120		
QY	411	CCTCCATATGTTACAATAGCTGGCCACCTAGAGGGTGTGATGTGCTGTGTSCAAACATGG	470		

121	CTCC-CATAGCTGCAATAGCTGGACACCCAGAGGCTGCAGAAGTCTGCTGCACATCG	179
471	AGCTAATGTCATAATTCAAGATGCAGTTTTTTCACTCCATTCGATATTGCAGCGTACTA	530
180	GGCCAAACGTGAATGTTCAAGATGCCGTCTCTTCACCCACTGCACATTCGACGCTACTA	239
531	TGGACATGAACAGAGTTAACTCGCCTCTCTTTTGAATTTGGTGGCTGAATGTAAGTCGG	590
240	TGGGCACGACGACGTTACCACTGTCCTTTTGAAGTTTGGTGTGATGTCATGTAAGCGG	299
591	TGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATCTTTGATATTCGAA	650
300	TGAAGTTGGGACACGGCCTCTGCACCTGGCCTCTGCAAGGGGCTCTTCAACATTTGAA	359
651	ACTCTTGATGGAAGAAGGACGACAAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGT	710
360	ACTCCTGGTAGAAGAGGAGCAAGCAGATGTGAACGCTCAGGACAAATGAAGACCACT	419
711	CCCACTCAATTTCTGTTCTCGAATTTGACACCAATGATATAGTTTAAGTATCTGCTGCAAG	770
420	CGCTCTGCACTTCTGTTCTCGAATTTGACACCAATATAGTAGGCTACTGCTCCAGAG	479
771	TGATTTGGAAGTTCAACCTCATGTTTGTTAATATCTATGGAGATACCCCTTACACCTGCG	830
480	TGACTTTAGAGGTCCAGCCTCAGCTCAATTAACATCTATGTTGACACTCTTTGCACTGCG	539
831	ATGCTCAATGGCAAAATTTGAAGTTGCCAGGAAATCATCCCAATATCAGGAACACAAAG	890
540	ATGCTACAAATGGAATTTTGAAGTTTGCAGGAAATTTGTCAGGTAAACAGAACTGAAG	599
891	TCTGACTAAGGAAACA	907
600	TCTGACTAAGGAAACA	616

RESULT 12

```

RESULT 12
US-09-833-381-1150
; Sequence 1150, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic
; FILE REFERENCE: 580C-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1150
; LENGTH: 736
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(736)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1150

```

	Query Match	16.6%	Score 414.6	DB 4	Length 736
	Best Local Similarity	81.2%	Pred. No. 3.8e-120		
	Matches 480	Conservative 0	Mismatches 111	Indels 0	Gaps 0
Qy	265	TTGAAGGGCTCCGGCCATCTCGACTCACAGAAGAAATGGATTTACAGCTTGCATTATGCA	324		
Db	1	TTAAAGGGCTCCGTCCATTCGAGACTCGAGAAATGGGTTTCCAGCTCTGGCACTGGCC	60		
Qy	325	GTTTACAAGGATAATGCAGAAATTCATCACTTCTCTGTTACAGTGGAGCTGATATACAG	384		
Db	61	GTTTACAGGACGCCCGGAATTATCACTTCACCTGTTGCACAGCGAGCATGTTTTCAG	120		
Qy	385	CAGGTGTGATAGGTGGGCTCACTGCCCTCCATATTGCTACAATAGCTGGGCACCTTAGAG	444		

```
Db 121 CAAAGTGGGATACGGTGGCTCCACAGCCCTCCACATAGCTCAATAGCTGGACACCCAGAG 180
QY 445 GCTGCTGATGCTGCTGTTGGAACATGAGAGCTAATGTCATATTAATCAAGATGCAATTTTTC 504
Db 181 GCTGAGAGTGTCTGCTACACATGGGCAATGTGAATGTTCAAGATGGCTGCTGCTTC 240
QY 505 ACTCCATTGCATATTGACAGCTACTATGACATCAAGAGGTAACCTCCCTTCTTTTGAAA 564
Db 241 ACCCACTGCATTTGAGCCCTACTATGGGACAGAGGTAACCAAGTTCCTTTTGAAG 300
QY 565 TTGTGCTGATGTAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCT 624
Db 301 TTGTGCTGATGTAATGTAAGTGGTGAAGTTGGGACAGCCCTCTGCACTGGCTCT 360
QY 625 GCAAAAGGATTTCTGAATATTGCAAACTCTTGATGGAGAGGACAGAGGATG 684
Db 361 GCAAAAGGCTTCTTCAACATTTGTGAACCTCTGTTAGAGAGGAGCAAGCAGATGT 420
QY 685 AATGCTCAAGATAATGAAGACCATGTCCCACTCCATTTCTGTTCTCGATTTGACACCAT 744
Db 421 AACGCTCAGGCAATGAAGACCATGCTCCCTCTGCACTTCTGTTCTCGATTTGACACCAT 480
QY 745 GATATAGTTAGTATCTCTGCTGCAAGTGAATTTGGAAGTTCAACTCATGTTGTTAATATC 804
Db 481 AATATAGTGAAGTACTCTCTCCAGAGTACTTANAGTCCAGGCTCAGCTCATTAACATC 540
QY 805 TATGAGATACCCCTTACACTGGCATGCTACAAATGCAAAATTTGAATTT 855
Db 541 TATGATGACATCTCTTGCACCTGGCATGCTACAAATTTGAATTT 591
```

RESULT 13

```
US-09-833-381-1151
; Sequence 1151, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1151
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(740)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1151
```

Query Match 15.5%; Score 413.2; DB 4; Length 740;

Best Local Similarity 79.0%; Pred. No. 1.1e-119; Mismatches 131; Indels 1; Gaps 1;

```
Matches 498; Conservative 0; Mismatches 131; Indels 1; Gaps 1;
QY 1 ATGGGAATTAATAATCTAGACCAACCAACTTGTTACTGATGAATGGAGAAAAGTC 60
Db 73 ATGGGAATTAATAATCTAGACCAACCAACTTGTTACTGATGAATGGAGAAAAGTC 132
QY 61 AGTGAATCATATGTTATCAATATGAAGATTAAGATCACTGAGATCAAGGAAA 120
Db 133 AGTGAATCTTACGCTATATCATAGAAGGCTGGAGATACTGAGATCAAGAAAAT 192
QY 121 GAACCTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAGTCAATTTA 180
Db 193 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCCTTCAAGTCAATTTA 252
QY 191 AATTACCGCACTGAATAAGGCTCTCTACTTCATTTATGTTGCAATTTTGTGGAGCAAG 240
```

```
Db 253 AATTACCGCAGAGCGTGGCTCTCCTGTACACCTCTGCTGTCTGTGGCGCAAC 312
QY 241 AATACATATTCGAACTCTTATGTTGAAAGGCTCCGCCATCTCGACTGACAGAAAT 300
Db 313 AAGTCATATCCGTGCCCTTATGTTAAAGGCTCCGTCCATCCAGACTGACGAAAT 372
QY 301 G3ATTTACAGCTTTGCAATTTAGCAGTTTACAAGGATAATGCAAAATGATCACTTCTCTG 360
Db 373 GGGTTTCAGCTCTGCACTGGCGTTTACAGGACAGCCCGGAATTTATCACTTCACTG 432
QY 361 CTTTAC-AGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCACTGCCCTCCATAT 419
Db 433 TTGCACAAAGGAGCAGATGTTTCAAGCAAGTGGGATACGGTGGCTCACAGCCCTCCACAT 492
QY 420 TGCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGTTGCTGTTGCAACATGGAGCTAATGT 479
Db 493 AGCTGCAATAGCTGGACACCCAGAGGCTGCANAAAGTCTGNTGCAACATGGGGCCAAAGT 552
QY 480 CAATATTTCAAGATGCACTTTTTTTTCACTCCATTCATATTCGAGCGTACTATGGACATGA 539
Db 553 GAATGTTCAAGATGCGCTCTTCTTCACTCCACTGCACATTCGAGCTACTATGGGACGA 612
QY 540 ACAGGTAACCTGCGCTCTTTTGAATTTGGTGTCTGATGTAATGTAAGTGTGAAGTTGG 599
Db 613 GCAGGTAACCAAGTGTCTCTTTTGAAGTTTGGNGGTGATGTCNATGTAAGCGGTAAAGTTGG 672
QY 600 AGATAGACCCCTCCACCTAGCATCTGCAAA 629
Db 673 GNAACAGGCTTCCGCCCTGCGCCNCGNAAA 702
```

RESULT 14

```
US-09-833-381-1147/c
; Sequence 1147, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1147
; LENGTH: 304
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(304)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1147
```

Query Match 7.4%; Score 184.2; DB 4; Length 304;

Best Local Similarity 79.8%; Pred. No. 8.6e-48; Mismatches 242; Conservative 0; Mismatches 59; Indels 3; Gaps 2;

```
QY 266 TGAAGGGCTCCGCCCATCTCGACTCACAAGAAATGGATTTACGCC--TTGCATTTAGC 323
Db 304 TAAAGGGCTCCGTCCATCCAGACTCAGAGAAATGGGTTTCCAAAGCTTTTGCACTGCC 245
QY 324 AGTTTACAGGATAATGCAAGATTTGATC-TCCTCTGCTTCACAGTGGAGCTGATATAC 382
Db 244 GGTTTACAAGGACAGCCCGGAATTTATCACTTCTACTGTTGACAGCGAGAGATGTC 185
QY 383 AGCAGGTTGGATACGGTGGCCCTCACTGCCCTCCATATTTGCTACAAATAGTGGCCACCTAG 442
Db 184 AGCAAGTGGGATACGGTGGCCCTCACAGCCCTCCACATAGCTGCAATAGCTGACACCCAG 125
QY 443 AGCTGCTGATGTTGCTTTGCAACATGGAGCTAAATGTCATATTCAGATGCGAGTTTTT 502
```

Db 124 AGGTCGAGAAGTCTGATGCAACATGGGGCCCAAGCTGAATGTTCAAGATCCGCTTCT 65
Qy 503 TCACTCCATTGCATATTGACGGTACTATGGACATGAACAGGTAACCTCTTTTGA 562
Db 64 TCACCCCACTGCACATGCGACNTACTATGGGACGAGGTAACCAAGTGTCTTTTGA 5
Qy 563 AATT 566
Db 4 AGTT 1

RESULT 15

US-09-833-381-1148/c
; Sequence 1148, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1148
; LENGTH: 186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(186)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1148

Query Match 3.4%; Score 84.4; DB 4; Length 186;
Best Local Similarity 69.1%; Pred. No. 1.7e-16;
Matches 112; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
Qy 405 CACTGCCCTCCATATTGCTACAAATAGCTGGCCACCTAGAGGCTGCTGCTGTTCCA 464
Db 162 CCCTTCCANNAATAGTGCCCAATAAGTTGNACACCCCAANAGGTGCAAAAGTCTGTTNCA 103
Qy 465 ACATGGAGCTAATGTCATATTCAAGATGAGTTTTTTTCACTCCATTCATATTGCAGC 524
Db 102 ACATGGGGCCCAATGTGAATTTCAAGATGCCGTCTTCTTCACCCACTGCACATTGCAGC 43
Qy 525 GTACTATGGACATGAACAGGTAACCTGCTTTCTTTTGAATT 566
Db 42 CTACTATGGGCACGAGCAGGTAACCAAGTGTCTTTTGAAGTT 1

Search completed: September 6, 2004, 07:42:53
Job time : 124.494 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2004, 23:38:08 ; Search time 4170.17 Seconds
(without alignments)
17938.068 Million cell updates/sec

Title: US-10-626-173-3

Perfect score: 2305

Sequence: 1 atgggaataataatctag.....gcagcttgagacagcgc 2505

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hrc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hrc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rtd:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	734.6	29.3	876	14	CF551822 AGENCOURT
2	721	28.8	756	12	BG227974 RST41894
3	682.8	27.3	813	12	BG209217 RST28732
4	677.4	27.0	793	12	BG216613 RST16305

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	665	26.5	665	13	BX501623	BX501623 DXF2p779A
6	618.6	24.7	902	10	BE897696	BE897696 601439226
7	587.4	23.4	656	13	BX502807	BX502807 DXF2p779C
8	557	22.2	638	9	AL599654	AL599654 DXF2p313B
9	491	19.6	491	13	BX112488	BX112488 BX112488
c 10	410.4	16.4	712	13	BQ622977	BQ622977 UI-H-Ful-
11	392.8	15.7	507	10	BF556094	BF556094 UI-R-Al-e
12	382.4	15.3	862	9	AU130402	AU130402 AU130402
13	374	14.9	950	14	CF412400	CF412400 CH3079 E
14	358.6	14.3	634	10	BB658710	BB658710 BB658710
15	355	14.2	465	13	BX099674	BX099674 BX099674
16	355	14.2	1113	14	CD513739	CD513739 AGENCOURT
c 17	347	13.9	684	9	AI375137	AI375137 EC09G12.x
c 18	345	13.8	648	13	BQ774566	BQ774566 UI-H-EZ1-
19	333.6	13.3	796	9	AU051569	AU051569 AU051569
c 20	330.6	13.2	514	10	BF197082	BF197082 7182b04.x
c 21	329.6	13.2	514	10	BF197116	BF197116 7183b04.x
22	324.4	13.0	667	14	CA589597	CA589597 hab36b06.
23	321.4	12.8	328	13	C03950	C03950 C03950 Huma
24	321.4	12.8	757	12	BG219957	BG219957 RST39729
25	321	12.8	820	13	B0313846	B0313846 603542120
c 26	317.4	12.7	530	9	AI333762	AI333762 qp98h10.x
27	310	12.4	446	12	BG406579	BG406579 dac14a04.
28	306.6	12.2	575	10	AW140835	AW140835 EST290917
29	301	12.0	623	10	B8659579	B8659579 BB559579
30	301	12.0	643	10	B8662065	B8662065 BB562065
31	301	12.0	705	11	AK052882	AK052882 Mus muscu
32	301	12.0	741	11	AK084817	AK084817 Mus muscu
33	286.8	11.4	673	10	BB480482	BB480482 BB480482
34	278	11.1	662	10	BB516995	BB516995 BB516995
c 35	269.2	10.7	574	13	B0447828	B0447828 UI-H-EU1-
36	259.8	10.4	693	12	BT547671	BT547671 603191815
37	255.2	10.2	699	13	B0358849	B0358849 603477541
c 38	254.2	10.1	519	10	BE501083	BE501083 7a36g06.x
39	250.8	10.0	535	14	CB718149	CB718149 AMGNNUC:N
40	247.6	9.9	834	13	BU386757	BU386757 603860892
41	238	9.5	531	14	CB719013	CB719013 AMGNNUC:N
c 42	237.4	9.5	452	9	AI991354	AI991354 ws09a11.x
c 43	232.2	9.3	413	9	AI040439	AI040439 owl5c03.x
c 44	225.2	9.0	384	13	BY086637	BY086637 BY086637
c 45	206.6	8.2	512	10	AW631150	AW631150 h161c06.x

ALIGNMENTS

RESULT 1
CF551822

LOCUS

DEFINITION

AGENCOURT 15595443 NIH MGC 183 Homo sapiens

IMAGE:30530169 5', mRNA sequence.

CF551822

CF551822

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 876)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabbs@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

CDNA Library Preparation: Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/ILNL at:

CF551822 AGENCOURT 15595443 NIH MGC 183 Homo sapiens linear EST 22-SEP-2003

IMAGE:30530169 5', mRNA sequence.

CF551822

CF551822

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 876)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabbs@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

CDNA Library Preparation: Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/ILNL at:

<http://image.llnl.gov>

```
ncsp: / / image: 1111:904
Plate: NPAM618 row: 0 column: 10
```

High quality sequence start: 15

High quality sequence	stop: 13
High quality sequence	stop: 670

FEATURES

source

1. 876

1. 876

L. 876
Organism="Homo sapiens"

```
/organs="HOME"
/mol time="mRNA"
```

```

/mol_type="mrna"
ref="taxon.0606"

```

/ab_xref="taxon:9606"
/ab_cno="IMAGE:30530169"
/ab_cno="IMAGE:30530169"

/clone="IMAGE:30530169"
/job boot="PHI08-T08X (T1 and the phase resistant)"

```
/lab_host="DHL0B-TONA"
/closo_1:b-INTU MCG 193
```

```
/clone lib="NIH MGC 183"
```

```
/note="Organ: pooled muscle (cardiac and skeletal);  
treat="Cover; cite=1; body (destroyed). Site 3.
```

Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:

NotI; Library is oligo-dT primed and directionally cloned (5'-GGG CCG GGA TTA CTC GAG AAT-3'). Average insert size is 1.8 kb.

(EcoRV site is destroyed upon cloning). Average insert

ORIGIN

Query Match	29.3%;	Score	734.6;	DB	14;	Length	876;
Best Local Similarity	97.6%;	Pred. No.	2.4e-189;				
Matches	830;	Conservative	0;	Mismatches	10;	Indels	10;
Gaps	8;						
QY	352	ACTTCTCTGCTTACAGCTGAGCTGATATACAGCAGGTTGGATACGGTGGCTCAGTGCC	411				
DB	28	ACTTCTCTGCTTC-CAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCAGTGCC	86				
QY	412	CTCCATATTTGCTACAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTGGCAATGGA	471				
DB	87	CTCCATATTTGCTACAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTGTGGCAATGGA	146				
QY	472	GCTAATGTCATATTTCAAGATGCAATTTTTTTTTCACCTCCATTCGATATTTGGAGCGTACTAT	531				
DB	147	GCTAATGTCATATTTCAAGATGCAATTTTTTTTTCACCTCCATTCGATATTTGGAGCGTACTAT	206				
QY	532	GGACATGAACAGGTAACCTCGCGCTCTTTTGAAATTTGGTGGCTGATGTAATTAAGTGGT	591				
DB	207	GGACATGAACAGGTAACCTCGCGCTCTTTTGAAATTTGGTGGCTGATGTAATTAAGTGGT	266				
QY	592	GAAGTTGGAGATAGACCCCTCCACTAGCATCTTGAAAAAGGATTCCTTGAATATTGCAAAA	651				
DB	267	GAAGTTGGAGATAGACCCCTCCACTAGCATCTTGAAAAAGGATTCCTTGAATATTGCAAAA	326				
QY	652	CTCTTGATGGAGAGGACGACCAAGCAGAGTGAATGCTCAAGATAATGAAGACCATGTC	711				
DB	327	CTCTTGATGGAGAGGACGACCAAGCAGAGTGAATGCTCAAGATAATGAAGACCATGTC	386				
QY	712	CCACTCCATTTCTGTTCTTCGATTTGGACACCATGATATAGTTAAAGTATCTGCTCAAAAGT	771				
DB	387	CCACTCCATTTCTGTTCTTCGATTTGGACACCATGATATAGTTAAAGTATCTGCTCAAAAGT	446				
QY	772	GAATTGGAAGTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCA	831				
DB	447	GAATTGGAAGTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCA	506				
QY	832	TGCTACAATGCAAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGT	891				
DB	507	TGCTACAATGCAAAATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGT	566				
QY	892	CTGACTAAGGAAAACATCTTCAGTGAACACGCTTTTCATAGTGCTTTGATACCTATGGCAAG	951				
DB	567	CTGACTAAGGAAAACATCTTCAGTGAACACGCTTTTCATAGTGCTTTGATACCTATGGCAAG	626				
QY	952	AGCATTGACCTAGTCAAAATTTCTCTTGATCAGAAATGCTATAAAATCAACCAACCAAGG-	1010				
DB	627	AGCATTGACCTAGTCAAAATTTCTCTTGATCAGAAATGCTATAAAATCAACCAACCAAGG	686				
QY	1011	AAGGGATGGGCACACT-GGATTACACTCT-GCTTGTCTACACGGTCACATTGCG-CCGTGT	1067				
DB	687	AAGGGATGGGCACACTGGATTACACTCTGCTGTCTACACGGTCACATTGCGCTGCTGT	746				
QY	1068	TCAGTCTTTTATCTGGATTAATGGAGCTGATATGAATCTTAGT-GGCTTGTGATCCAGCAAG-	1125				

Db	747	TCAGTCTTCTGATGAATCGAGCTGATATGAATCTAGTGGCTTGTCATCCAGCAGGN	806			
Qy	1126	TCATAGTGGTGAAGAGATGAGCAGACATGTTT--CATGTGGGCTTATGAAAAAGGCGCAT	1182			
Db	807	TCATAGTGGTGAAGAGATGAGCAGACATGTTTGTAGTGGGCTTATGAAAAAGGCGCAT	866			
Qy	1183	GATGCCATTG	1192			
Db	867	GATGCCATTG	876			
RESULT 2						
BG227974						
LOCUS	BG227974 756 bp mRNA linear EST 21-APR-2001					
DEFINITION	RST41894 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.					
ACCESSION	BG227974					
VERSION	BG227974.1 GI:13748097					
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
	1 (bases 1 to 756)					
	Harrington,J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,					
AUTHORS	Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,					
	Whittington, J., Lerner, L., Coscanzo, D., McElligott, K., Booser, S.,					
	Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,					
	Offenbacher, J., Danzig, J. and Ducar, M.					
Creation of genome-wide protein expression libraries using random						
activation of gene expression						
Nac. Biotechnol. 19 (5), 440-445 (2001)						
JOURNAL	21227151					
MEDLINE	11329013					
PUBMED	Contact: Scott J. Cain					
COMMENT	Athersys, Inc.					
FEATURES	3201 Carnegie Ave, Cleveland, OH 44115, USA					
	Tel: 216 431 9900					
	Fax: 216 361 9596					
	Email: scain@athersys.com					
High quality sequence stop: 520.						
source	Location/Qualifiers					
	1..756					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
ORIGIN	/db_xref="taxon:9606"					
	/cell_line="HT1080"					
	/clone_lib="Athersys RAGE Library"					
	/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', the Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."					
Query Match 28.8%; Score 721; DB 12; Length 756;						
Best Local Similarity 97.3%; P-Value 1.2e-185;						
Matches 733; Conservative 0; Mismatches 20; Indels 0; Gaps 0;						
Qy	546	AACTCGCCTCTTTTGGAAATTTGGTGCTCATGTAAATGTAAAGTGTGTAAGTTGGAGATAG	605			
Db	1	AACTCGCCTCTTTTGGAAATTTGGTGCTCATGTAAATGTAAAGTGTGTAAGTTGGAGATAG	60			
Qy	606	ACCCCTCACCCTAGCTCTCCAAAAGGATCTTGAATATGCAAACTCTTGATGGAAGA	665			
Db	61	ACCCCTCACCCTAGCTCTCCAAAAGGATCTTGAATATGCAAACTCTTGATGGAAGA	120			
Qy	666	AGGCAGCAAAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCATTTCTG	725			
Db	121	AGGCAGCAAAAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCATTTCTG	180			
Qy	726	TTTCGATTGGACACCATCATATAGTAAATCTGTGTGAAAGTGAATTTGGAAGTTCA	785			


```

LOCUS      BG216613              793 bp    mRNA    linear    EST 21-APR-2001
DEFINITION RST36305 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION  BG216613
VERSION    BG216613.1 GI:13742634
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 793)
AUTHORS   Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
          Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
          Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
          Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
          Offenbacher,J., Danzig,J. and Ducar,M.
          Creation of genome-wide protein expression libraries using random
          activation of gene expression
          Nat. Biotechnol. 19 (5), 440-445 (2001)
JOURNAL   Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE   21227151
PUBMED    11329013
COMMENT   Contact: Scott J. Cain
          Athersys, Inc.
          3201 Carnegie Ave, Cleveland, OH 44115, USA
          Tel: 216 431 9900
          Fax: 216 361 9596
          Email: scain@atersys.com
          High quality sequence stop: 511.
          Location/Qualifiers
            1..793
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone_lib="HT1080"
              /note="See 'Creation of Genome-wide Protein Expression
              Libraries using Random Activation of Gene Expression',
              Nature Biotechnology, in press. Note that even though the
              cell type indicated is HT1080, since a random activation
              method was used, these sequence tags are not necessarily
              expressed in HT1080 under normal circumstances."

ORIGIN
Query Match      27.0%; Score 677.4; DB 12; Length 793;
Best Local Similarity 96.4%; Pred. No. 1e-173;
Matches 767; Conservative 0; Mismatches 22; Indels 7; Gaps 7;

QY 546 AACTCGCCCTCTTTTGAATTTGGTGGTGGTGAATTAAGTGAAGTGGAGATAG 605
DB 1 AACTCGCCCTCTTTTGAATTTGGTGGTGGTGAATTAAGTGAAGTGGAGATAG 60

QY 606 ACCCTCCACCTAGCATCTGCAGAAAGGATTTCTTGAATATTGCAAAACTCTTGATGGAAG 665
DB 61 ACCCTCCACCTAGCATCTGCAGAAAGGATTTCTTGAATATTGCAAAACTCTTGATGGAAG 120

QY 666 AGCAGGAGAACGATGATGCTCAAGATATGAAGACCATGTCCTCCACTCCATTCTG 725
DB 121 AGCAGGAGAACGATGATGCTCAAGATATGAAGACCATGTCCTCCACTCCATTCTG 180

QY 726 TTCTCGATTGTCACACCATGATATAGTTAAGTATCTGCTCAAAGTGATTGGAAGTTCA 785
DB 181 TTCTCGATTGTCACCATGATATAGTTAAGTATCTGCTCAAAGTGATTGGAAGTTCA 240

QY 786 ACCTCATGTTTGAATATCTATGAGATACCCCTTTACACCTGGCATGCTACAAATGGCAA 845
DB 241 ACCTCATGTTTGAATATCTATGAGATACCCCTTTACACCTGGCATGCTACAAATGGCAA 300

QY 846 ATTTGAAGTTCCCAAGGAATATCTCCAAATATCAGGAACAGAAGTCTCACTAAGGAAA 905
DB 301 ATTTGAAGTTCCCAAGGAATATCTCCAAATATCAGGAACAGAAGTCTCACTAAGGAAA 360

QY 906 CATCTTCAGTGAACAGCTTTTCATAGTGTCTGTACCTATGCGAAGAGCATTCAGCTAGT 965
DB 361 CATCTTCAGTGAACAGCTTTTCATAGTGTCTGTACCTATGCGAAGAGCATTCAGCTAGT 420

```

```

QY 966 CAAATTTCTTTGATCAGAAATCTATAAACATCAACCAAGGAGGATGGGCACAC 1025
DB 421 CAAATTTCTTTGATCAGAAATCTATAAACATCAACCAAGGAGGATGGGCACAC 480

QY 1026 TGGATTACACTCTGCTTACCAACGGTCACATTCGCTGGTTCAGTTCTTACTCGATAA 1085
DB 481 TGGATTACACTCTGCTTACCAACGGTCACATTCGCTGGTTCAGTTCTTACTCGATAA 540

QY 1086 TGGAGCTGATATGAATCTAGTGGCTTGTGATCCACAGGCTAGTGGTGAAGAAAGATCA 1145
DB 541 TGGAGCTGATATGAATCTAGTGGCTTGTGATCCACAGGCTAGTGGTGAAGAAAGATCA 600

QY 1146 GCAGACATGTTGATGTTGGCTTATGAAAAGGCGCATGATGCGATTGTCACACTCTGAA 1205
DB 601 CCAGACATGTTGATGTTGGCTTATGAAAAA-GGCGTGATGCTTTGTC-CACTCTCTGAA 658

QY 1206 GCATTATAAGAGACCAAGATGAATTGCCCTGTGAATGAATATTTCTCAGCTTGGAGGAGA 1265
DB 659 GC-TTATTAAGAGACC-CCAGATGAATTGCCCTG-HATGAATATTTCTCACCTT-GAAGAGA 714

QY 1266 TGGCTCCTATGTGTCTTCCATCACCCTTGGGGAAGATTAAGCATGACAAAAAGAGAA 1325
DB 715 TGGCTCCTATGTGTCTTCCATCACCCTTGGGGA-ATTAAGAGCTTGACCCAGAGAG 773

QY 1326 GGCAGATATTCCTCCTC 1341
DB 774 GGCAGATATTCCTCCTC 789

```

```

RESULT 5
LOCUS    BX501623              665 bp    mRNA    linear    EST 04-SEP-2003
DEFINITION DKFZp779A1164.x1.779 (synonym: hnccl) Homo sapiens cDNA clone
ACCESSION BX501623
VERSION    BX501623
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

```

```

REFERENCE 1 (bases 1 to 665)
AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
          Fobo,G., Han,M. and Wiemann,S.
          EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
          Unpublished (2003)
          Contact: MIPS
          MIPS

```

```

JOURNAL   Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
COMMENT    This is the 5' sequence of the clone insert
          Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
          Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
          sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
          consortium of the German Genome Project.
          No 5' sequence available.
          This clone (DKFZp779A1164) is available at the RZPD in Berlin.
          Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
          Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```

```

FEATURES             source
     1..665
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="DKFZp779A1164"
         /tissue_type="liver"
         /dev_stage="fetal"
         /lab_host="DH108"
         /clone_lib="779 (synonym: hnccl)"
         /note="Vector: pSport1_Sfi; Site_1: SfiA; Site_2: SfiB"

```

```

ORIGIN
Query Match      26.5%; Score 665; DB 13; Length 665;

```

Best Local Similarity 100.0%; Pred. No. 2.3e-170;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 526 TACTATGGACATGAACAGCTAACTGCGCTCTTTTGAATTTGGTGGCTGATGAATGTA 585
Db 1 TACTATGGACATGAACAGCTAACTGCGCTCTTTTGAATTTGGTGGCTGATGAATGTA 60

QY 586 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAATATT 645
Db 61 AGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTGAATATT 120

QY 646 GCAAACTCTTGATGGAAGAGGAGAGAGAGAGATGGAATGCTCAAGATAAGAGAC 705
Db 121 GCAAACTCTTGATGGAAGAGGAGAGAGAGAGATGGAATGCTCAAGATAAGAGAC 180

QY 706 CATGTGCCACTCCATTTCTGTTCTGATTTGACACCATGATATAGTTAAAGTATCTGCTG 765
Db 181 CATGTGCCACTCCATTTCTGTTCTGATTTGACACCATGATATAGTTAAAGTATCTGCTG 240

QY 766 CAAAGTATTTGGAAGTTCAACCTCATGTTGTTATATCTATGGAGATACCCCTTACAC 825
Db 241 CAAAGTATTTGGAAGTTCAACCTCATGTTGTTATATCTATGGAGATACCCCTTACAC 300

QY 826 CTGGCATGCTACAAATGGAAGTTGCAAGGAAATCATCAAAATATCAGGAACA 885
Db 301 CTGGCATGCTACAAATGGAAGTTGCAAGGAAATCATCAAAATATCAGGAACA 360

QY 886 GAAAGTCTGACTAAGGAAACATCTTTCACTGAAACAGCTTTTTCATAGTCTTTACCTAT 945
Db 361 GAAAGTCTGACTAAGGAAACATCTTTCACTGAAACAGCTTTTTCATAGTCTTTACCTAT 420

QY 946 GGAAGAGCATTTGACCTAGTCAAAATTTCTTCTGATCAGAAATGTCATAAATCAACAC 1005
Db 421 GGAAGAGCATTTGACCTAGTCAAAATTTCTTCTGATCAGAAATGTCATAAATCAACAC 480

QY 1006 CAAAGGAGGATGGGACACCTGATTTACACTCTGCTTGTACACGGTCACTTGGCCTG 1065
Db 481 CAAAGGAGGATGGGACACCTGATTTACACTCTGCTTGTACACGGTCACTTGGCCTG 540

QY 1066 GTTCAGTCTTACTGATATGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGG 1125
Db 541 GTTCAGTCTTACTGATATGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGG 600

QY 1126 TCTAGTGGTGAAGATGAGCAGACATGTTTGTAGTGGCTTATGAAGAGGCGATCAT 1185
Db 601 TCTAGTGGTGAAGATGAGCAGACATGTTTGTAGTGGCTTATGAAGAGGCGATCAT 660

QY 1186 GCCAT 1190
Db 661 GCCAT 665

RESULT 6
BE897696
LOCUS 601439226F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3924055 5',
DEFINITION mRNA sequence.
ACCESSION BE897696
VERSION BE897696.1 GI:10363419
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 902)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>,
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/dmp
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLN9761 row: d column: 08
High quality sequence stop: 689.
Location/Qualifiers
1. .902
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3924055"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_72"
/note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 24.7%; Score 618.6; DB 10; Length 902;
Best Local Similarity 96.2%; Pred. No. 1.3e-157;
Matches 678; Conservative 0; Mismatches 19; Indels 8; Gaps 4;

QY 332 AGGATAATGCAGAAATGATCACTTCTGCTTCACAGTGGAGCTGATATACAGAGGTTG 391
Db 44 AGGATAATGCAGAAATGATCACTTCTGCTTCACAGTGGAGCTGATATACAGAGGTTG 103

QY 392 GATACGGTGGCTCACTGCGCTCCCATATTGCTACAAATAGCTGGCCACCTAGAGGCTGCTG 451
Db 104 GATACGGTGGCTCACTGCGCTCCCATATTGCTACAAATAGCTGGCCACCTAGAGGCTGCTG 163

QY 452 ATGTGCTGTGCAACATGAGAGCTAATGCAATATTCAGAGTGCAGTCTTTTCACTCCAT 511
Db 164 ATGTGCTGTGCAACATGAGAGCTAATGCAATATTCAGAGTGCAGTCTTTTCACTCCAT 223

QY 512 TGCAATATTCAGCGTACTATGACATGAACAGTAACCTCGCCTCTCTTTGAAATTTGGTG 571
Db 224 TGCAATATTCAGCGTACTATGACATGAACAGTAACCTCGCCTCTCTTTGAAATTTGGTG 283

QY 572 CTGATCTAAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAG 631
Db 284 CTGATCTAAATGTAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAG 343

QY 632 GATTCCTGAATATTCGAAACCTCTTGTGGAAGAGGCGAGCAAGCAGATGTGAATGCTC 691
Db 344 GATTCCTGAATATTCGAAACCTCTTGTGGAAGAGGCGAGCAAGCAGATGTGAATGCTC 403

QY 692 AAGATAATGAAGACCATGTCCCACTCCCATTTCTGTTCTCGAATTTGGACACCATGATATAG 751
Db 404 AAGATAATGAAGACCATGTCCCACTCCCATTTCTGTTCTCGAATTTGGACACCATGATATAG 463

QY 752 TTAAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACCTCATGTTGTTAATCTATGAG 811
Db 464 TTAAGTATCTGCTGCAAGTGAATTTGGAAGTTCAACCTCATGTTGTTAATCTATGAG 523

QY 812 ATACCCCTTACACCTGGCATGCTACAATGGCAAAATTTGAAGTTGCCAAGGAAATCATCC 871
Db 524 ATACCCCTTACACCTGGCATGCTACAATGGCAAAATTTGAAGTTGCCAAGGAAATCATCC 582

QY 872 AAATATCAGAAACAGAAAGTCTGACTAGAGAAAACATCTTCTGAGTGAACACGCTTTTCATA 931
Db 583 AAATATCAGAAACAGAAAGTCTGACTGACTAGAGAAAACATCTTCTGAGTGAACACGCTTTTCATA 641

QY 932 GTGCTGTACCTATGCGAAGAGCATTTGACCTAGTCAAAATTTCTTCTGATCAGAAATGTCA 991
Db 642 GTGCTGTACCTATGCGAAGAGCATTTGACCTAGTCAAAATTTCTTCTGATCAGAAATGTCT 698

QY 992 TAAACATCAACCAAGGAGGATGGGCACACTGGATTACAT 1036
Db 699 TAGACATTCACACACAGGAA---GGATGGCACACTGGATTACTCT 740

```

RESULT 7
BX502807      656 bp      mRNA      linear      EST 04-SEP-2003
LOCUS        DKFZp779C2070 r1 779 (synonym: hnccl) Homo sapiens cDNA clone
DEFINITION   DKFZp779C2070_5', mRNA sequence.
ACCESSION   BX502807
VERSION     BX502807.1 GI:32024614
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 656)
AUTHORS     Bahr,A., Lauber,J., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
TITLE       EST (Bahr,A., Lauber,J., Mewes,H.W., Weil,B., et al.)
JOURNAL     Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No s1 sequence available.
            This clone (DKFZp779C2070) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        1..656
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZp779C2070"
            /tissue_type="liver"
            /dev_stage="fetal"
            /lab_host="DH10B"
            /clone_lib="779 (synonym: hnccl)"
            /note="Vector: pSpout1_Sfi; Site_1: SfiI; Site_2: SfiIB"
ORIGIN
    Query Match      23.4%; Score 587.4; DB 13; Length 656;
    Best Local Similarity 99.7%; Pred. No. 4e-149;
    Matches 588; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGGAATTATTAATCTAGACCAACCCAACTTGCTACTGATGAATGGAGAAAAGTC 60
DB 67 ATGGGAATTATTAATCTAGACCAACCCAACTTGCTACTGATGAATGGAGAAAAGTC 126
QY 61 AGTGAATCATATGTTATCACAAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120
DB 127 AGTGAATCATATGTTATCACAAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 186
QY 121 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAAGCCCTTCAGTAAGTCAATTTA 180
DB 187 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAAGCCCTTCAGTAAGTCAATTTA 246
QY 181 AATTACCCGACTGAAATGGGCTGCTCTACTTCATTTATGTTGCAATTTGTGGAGCAAG 240
DB 247 AATTACCCGACTGAAATGGGCTGCTCTACTTCATTTATGTTGCAATTTGTGGAGCAAG 306
QY 241 AAATCACATATTCGAATCTTTATGTTGAAAGGGCTCGCCCATCTCGATGACAGAAAT 300
DB 307 AAATCACATATTCGAATCTTTATGTTGAAAGGGCTCGCCCATCTCGATGACAGAAAT 366
QY 301 GGATTTACAGCTTGCATTTAGCAATTTACAGGATAATGCAAGTTGATCACTTCTCTG 360
DB 367 GGATTTACAGCTTGCATTTAGCAATTTACAGGATAATGCAAGTTGATCACTTCTCTG 426
QY 361 CTTACAGTGGAGCTGTATATACAGAGTTGGATACGGTGGCTCTACCTCCCTCCATATT 420

```

```

DB 427 CTTACAGTGGAGCTGTATATACAGAGTTGGATACGGTGGCTCTACTCCCTCCATATT 486
QY 421 GCTACAATAGCTGCCACCTAGAGGTGCTGATGTCTCTTTGCAACATGAGAGCTAATGTC 480
DB 487 GCTACAATAGCTGCCACCTAGAGGTGCTGATGTCTCTTTGCAACATGAGAGCTAATGTC 546
QY 481 AATATTCAAGATGAGCTTTTTCACCTCCATTTGCAATTTGCGAGCTTACTATGACATGAA 540
DB 547 AATATTCAAGATGAGCTTTTTCACCTCCATTTGCAATTTGCGAGCTTACTATGACATGAG 606
QY 541 CAGGTAATCTGCCTCTTTTGAATTTGGTGTGCTGATGATAAATGTAAGTGG 590
DB 607 CAGGTAATCTGCCTCTTTTGAATTTGGTGTGCTGATGATAAATGTAAGTGG 656

AL599654      638 bp      mRNA      linear      EST 04-SEP-2003
LOCUS        DKFZp313B0430 r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
DEFINITION   DKFZp313B0430_5', mRNA sequence.
ACCESSION   AL599654
VERSION     AL599654.1 GI:15162942
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 638)
AUTHORS     Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.
TITLE       EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL     Unpublished (1999)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            No s1 sequence available.
            This clone (DKFZp313B0430) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        1..638
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZp313B0430"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="313 (synonym: hlcc2)"
            /note="Vector: pTriplex2; Site_1: SfiI; Site_2: SfiIB;
            cDNA-collection"
ORIGIN
    Query Match      22.2%; Score 557; DB 9; Length 638;
    Best Local Similarity 100.0%; Pred. No. 8.3e-141;
    Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGAATTATTAATCTAGACCAACCCAACTTGCTACTGATGAATGGAGAAAAGTC 60
DB 82 ATGGGAATTATTAATCTAGACCAACCCAACTTGCTACTGATGAATGGAGAAAAGTC 141
QY 61 AGTGAATCATATGTTATCACAAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120
DB 142 AGTGAATCATATGTTATCACAAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 201
QY 121 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAGCCCTTCAGTAAGTCAATTTA 180
DB 202 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAGCCCTTCAGTAAGTCAATTTA 261

```

QY 181 AATTACCGCAGTGAATGGCTGCTCTACTTCAATTTATGTCATTTGGGAGCAAG 240
Db 262 AATTACCGCAGTGAATGGCTGCTCTACTTCAATTTATGTCATTTGGGAGCAAG 321
QY 241 AATACATATTGCAACTCTTATGTTGAAGGCTCCGCCCATCTCGACTGACAGAAAT 300
Db 322 AATACATATTGCAACTCTTATGTTGAAGGCTCCGCCCATCTCGACTGACAGAAAT 381
QY 301 GGNATTACAGCCTTGCAATTTACAGCTTTACAGGATGATGAGAAATGACATCTCTCTG 360
Db 382 GGNATTACAGCCTTGCAATTTACAGGATGATGAGAAATGACATCTCTCTG 441
QY 361 CTTACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCAGTCGCCCTCCATATT 420
Db 442 CTTACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCAGTCGCCCTCCATATT 501
QY 421 GTCACATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGCGCAACATGAGAGCTAATGTC 480
Db 502 GTCACATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGCGCAACATGAGAGCTAATGTC 561
QY 481 AATATTCAAGATGCACTTTTTCATCTCCATGTCATATTGAGCGTACTATGGACATGAA 540
Db 562 AATATTCAAGATGCACTTTTTCATCTCCATGTCATATTGAGCGTACTATGGACATGAA 621
QY 541 CAGGTAACCTGGCTTCT 557
Db 622 CAGGTAACCTGGCTTCT 638

RESULT 9
BX112488
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

112488 Soares fetal lung NbHL19W Homo sapiens cDNA clone
IMAGE:1931107, mRNA sequence.
BX112488.1 GI:27878929
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
Human Unigeneset - R2PD3
Unpublished (2003)
Contact: Ina Rols
R2PD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
R2PD; IMAGp998D204738.
R2PDLiB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - R2PD3 (R2PDLiB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rols
R2PD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de

FEATURES
source
1. 491
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998D204738 ; IMAGE:1931107"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_lung_NbHL19W"
/notes="Organ: lung; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo (dT) primer
[5'-TGTTACCAATCTAAGTGGAGCGCGCAATTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NbHL19W."

ORIGIN
Query Match 19.6%; Score 491; DB 13; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.5e-123;
Matches 491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1109 CTTGTGATCCCGAGCGCTAGTGGTGAAGAAGATGACACATGTTGATGGGCTT 1168
Db 1 CTTGTGATCCCGAGCGCTAGTGGTGAAGAAGATGACACATGTTGATGGGCTT 60
QY 1169 ATGAAAGAGGCGATGATGCCATTGTGCACATCTCTGAGCATTTAAGAGAGCACAAGATG 1228
Db 61 ATGAAAGAGGCGATGATGCCATTGTGCACATCTCTGAGCATTTAAGAGAGCACAAGATG 120
QY 1229 AATTGCCCTGTAATGAATATTTCTAGCTGGAGAGATGGCTCTATGTGCTTCCAT 1288
Db 121 AATTGCCCTGTAATGAATATTTCTAGCTGGAGAGATGGCTCTATGTGCTTCCAT 180
QY 1289 CACCTTTGGGGAAGATTAAAGCATGACAAAGAGAGAGAGATATTCCTCTCTAAGAG 1348
Db 181 CACCTTTGGGGAAGATTAAAGCATGACAAAGAGAGAGAGATATTCCTCTCTAAGAG 240
QY 1349 CTGGATTGCTTCAATTTCCATTTCCATCTCAGCTCTCAGAAATTTAGTTCATGAGATTATTG 1408
Db 241 CTGGATTGCTTCAATTTCCATTTCCATCTCAGCTCTCAGAAATTTAGTTCATGAGATTATTG 300
QY 1409 GCTCAGGTTCTTTTGGGAAGTATATAAGGACGATGAGAAATAAATACTGGCTATAA 1468
Db 301 GCTCAGGTTCTTTTGGGAAGTATATAAGGACGATGAGAAATAAATACTGGCTATAA 360
QY 1469 AAGTTATCGAGCCAAATACCTACTCTCCAGTCAGATGTGGATATGTTTCCGAGAGG 1528
Db 361 AAGTTATCGAGCCAAATACCTACTCTCCAGTCAGATGTGGATATGTTTCCGAGAGG 420
QY 1529 TGTCCATTTCTGCCAGCTCAATCATCTCCGTAATTTAGTTCGCTGCTTGA 1588
Db 421 TGTCCATTTCTGCCAGCTCAATCATCTCCGTAATTTAGTTCGCTGCTTGA 480
QY 1589 ATGATCCCGAGC 1599
Db 481 ATGATCCCGAGC 491

RESULT 10
BU622977/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

10 BU622977.1 GI:23289192
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsof@mail.nih.gov
Tissue Procurement: James Martin

712 bp mRNA linear EST 23-SEP-2002
UI-H-FL1-bgb-n-08-0-UI.s1 NCI CGAP FL1 Homo sapiens cDNA clone
UI-H-FL1-bgb-n-08-0-UI.3, mRNA sequence.

BU622977
UI-H-FL1-bgb-n-08-0-UI.s1 NCI CGAP FL1 Homo sapiens cDNA clone
UI-H-FL1-bgb-n-08-0-UI.3, mRNA sequence.

BU622977.1 GI:23289192
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsof@mail.nih.gov
Tissue Procurement: James Martin

1762 GACTTGAAAGTCACAAATATCTTCTATATAGGATGGGCATGCTGTGGTGGAGATTTT 1821
|||
68 GACCTGAAAGCAGCACAATATCTGTCTATGAGGATGCCATGCTGTGGTGGAGATTTT 127
|||
1822 GGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAACATGACAAACACCTGGGAAC 1881
|||
128 GGAGATCAAGATTTCTGAGTCCCTGGATGAAGACAACATGACAAACACCTGGGAAC 187
|||
1882 CTCGGTGGATGGCTCTCTGAGTGTTCACGCGAGTGCATCTCGGTACACCATCAAGACGAT 1941
|||
188 CTCGGTGGATGGCTCTGAGTGTTCACACAGTGCACGAGATACACCATCAAGCTGAT 247
|||
1942 GCTTCAGTATGCTGTGTCTGTGGGAAATCTCACTGGGAAATTCGATTCGCTCAT 2001
|||
248 GCTTCAGTATGCTGTGTCTGTGGGAGCTCCCTCACTGGGAAATTCGATTCGCTCAT 307
|||
2002 CTCAGACCCAGCGCTCGGCGACGACATGCTTACCACCATCAGACCTCCCATTTGGC 2061
|||
308 CTCAGACCCAGCGCTCGGCGACGACATGCTTACCACCATCAGACCTCCCATCGGC 367
|||
2062 TATTCATTTCCAGCCCATATCATCTCTGTGTGATACGAGGTGGAACGATGCTCGAA 2121
|||
368 TATTCATTTCCAGCCCATATCATCTCTGTGTGATACGAGGTGGAATGCTCGTGA 427
|||
2122 GGAGACCCGAAATTTCTGAAGTTGTCATGAAGTTAGAAAGTGTCTCTGCAACATTGAG 2181
|||
428 GGAGACCCGAAATTTCTGAAGTTGTCATGAAGTTAGAAAGTGTCTCTGCAACATTGAG 487
|||
2182 CTGATGCTCTCGCATCAAG 2201
|||
488 CTCATGCTCCAGCATCAAG 507
|||

RESULT 12
LOCUS AU130402 862 bp mRNA linear EST 01-AUG-2002
DEFINITION AU130402 NT2RP3 Homo sapiens cDNA clone NT2RP3000778 5', mRNA
sequence.

ACCESSION AU130402
VERSION AU130402.1 GI:10990756
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 862)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL HRI human cDNA project
COMMENT Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES
source
1. .862
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3000778"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RP3"
/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor

ORIGIN
Query Match 15.3%; Score 382.4; DB 9; Length 862;
Best Local Similarity 99.0%; Pred. No. 5.8e-93;
Matches 394; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 38 CTGATGAATGGAGAAAAAGTCAGTGAATCATATGTTATCACAATAGAAAGATTGAAAG 97
|||
Db 376 CTGATGAATGGAGAAAAAGTCAGTGAATCATATGTTATCACAATAGAAAGATTGAAAG 435
|||
QY 98 ATGACCTGCGAGATCAAGAAAAAGAACTGCAGACAACCTAAAGGAATATATTTGGCTCTCATG 157
|||
Db 436 ATGACCTGCGAGATCAAGGAAAAAGAACTGCAGACAACCTAAAGGAATATATTTGGCTCTCATG 495
|||
QY 158 AAGCCTTCAGTAAAGTCATTTAAATTTACCGACTGAAAAATGGGCTGTCTTACTTCAAT 217
|||
Db 496 AAGCCTTCAGTAAAGTCATTTAAATTTACCGACTGAAAAATGGGCTGTCTTACTTCAAT 555
|||
QY 218 TATGTTGCATTTGTGGAGCAAGAAATCACATATTCGAACTCTTATGTTGAAAGGGCTCC 277
|||
Db 556 TATGTTGCATTTGTGGAGCAAGAAATCACATATTCGAACTCTTATGTTGAAAGGGCTCC 615
|||
QY 278 GCCATCTCGACTGACAAGAAATGGAATTTACAGCTTGCATTTAGCAGTTTACAGGATA 337
|||
Db 616 GCCATCTCGACTGACAAGAAATGGAATTTACAGCTTGCATTTAGCAGTTTACAGGATA 675
|||
QY 338 ATGCAGAAATTTGATCACTTCTCTGCTTACAGTGGAGCTGATATACAGCAGTTTGGATACG 397
|||
Db 676 ATGCAGAAATTTGATCACTTCTCTGCTTACAGTGGAGCTGATATACAGCAGTTTGGATACN 735
|||
QY 398 GTGGCCTCACTGCCCTCCATATTTGCTACAATAGCTGGC 435
|||
Db 736 GTGGCCTCACTG-CCTTCATATTTGCTACAATAGCTGGC 772
|||

RESULT 13
LOCUS CF412400 950 bp mRNA linear EST 02-SEP-2003
DEFINITION CF412400 E08MF Canine heart normalized cDNA Library in paluascript
Canis familiaris cDNA clone CH3#079_E08 5', mRNA sequence.

ACCESSION CF412400
VERSION CF412400.1 GI:34413646
KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 950)
AUTHORS Yi,Y., Desai,R., Olarte,M., Henthorn,P. and George A.L.
TITLE Expressed sequence tags from Canine heart
JOURNAL Unpublished (2003)
COMMENT Other ESTs: CH3#079_E08MR
Contact: George Al.
Division of Genetic Medicine
Vanderbilt University
529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA
Tel: 615 936 2660
Fax: 615 936 2661
Email: al.george@vanderbilt.edu
Insert Length: 1883 Std Error: 0.00
Seq primer: MF: GTTTTCCAGTCACGACGTTG
High quality sequence start: 68
High quality sequence stop: 441.
High quality sequence stop: 441.
Location/Qualifiers
1. .950
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/clone="CH3#079_E08"
/tissue_type="heart"
/cell_type="heart"
/dev_stage="mixed developmental stages (adult, 30 day - 40

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 2, 2004, 14:15:41 ; Search time 136 Seconds
(without alignments)
1937.189 Million cell updates/sec

Title: US-10-626-173-2

Perfect score: 4390

Sequence: 1 MGNYKSRPTCTDEWKKV.....PMSSVHFHSCRNSSEFEDSS 835

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4390	100.0	835	4 Q9Y2V6	Q9Y2V6 homo sapien
2	4079	92.9	835	11 Q7TQP6	Q7TQP6 rattus norv
3	1716.5	39.1	850	5 O17055	O17055 caenorhabdi
4	672	15.3	165	11 Q8BFR0	Q8BFR0 mus muscucu
5	486	11.1	546	10 O22558	O22558 arabidopsis
6	482	11.0	603	10 Q7XXN2	Q7XXN2 oryza sativ
7	476.5	10.9	445	10 Q9LWF8	Q9LWF8 arabidopsis
8	476	10.8	525	10 Q9FIL6	Q9FIL6 arabidopsis
9	475.5	10.8	470	10 Q7X153	Q7X153 oryza sativ
10	475	10.8	438	10 Q9X187	Q9X187 arabidopsis
11	475	10.8	438	10 Q89J9	Q89J9 arabidopsis
12	473.5	10.8	553	10 Q81808	Q81808 arabidopsis
13	472	10.8	525	10 Q8WCZ2	Q8WCZ2 arabidopsis
14	469.5	10.7	570	10 Q8RWL6	Q8RWL6 arabidopsis
15	469.5	10.7	638	5 Q86AT8	Q86AT8 dictyosteli
16	460.5	10.5	1719	4 Q13768	Q13768 homo sapien

17	454	10.3	1856	4	Q99407	Q99407 homo sapien
18	446	10.2	1549	5	Q9V4B1	Q9V4B1 drosophila
19	444	10.1	1136	6	Q9N180	Q9N180 bos taurus
20	444	10.1	1848	11	Q61302	Q61302 mus muscucu
21	443	10.1	1549	5	Q24241	Q24241 drosophila
22	442	10.1	422	10	Q8GV29	Q8GV29 oryza sativ
23	442	10.1	422	10	Q7XPE4	Q7XPE4 oryza sativ
24	442	10.1	743	5	Q9Y1V2	Q9Y1V2 ephydatia f
25	439.5	10.0	482	10	Q39886	Q39886 glycine max
26	438.5	10.0	417	10	Q8GV30	Q8GV30 oryza sativ
27	438.5	10.0	616	13	Q73613	Q73613 xenopus lae
28	438	10.0	412	10	Q9M085	Q9M085 arabidopsis
29	436.5	9.9	919	4	Q8N8A2	Q8N8A2 homo sapien
30	436	9.9	411	10	Q9ZQ31	Q9ZQ31 arabidopsis
31	435.5	9.9	638	4	Q81272	Q81272 homo sapien
32	435	9.9	1863	4	Q7Z3L5	Q7Z3L5 homo sapien
33	434.5	9.9	541	10	Q7XSP3	Q7XSP3 oryza sativ
34	431.5	9.8	1943	11	Q61307	Q61307 mus muscucu
35	430	9.8	1219	11	Q8C8R3	Q8C8R3 mus muscucu
36	428.5	9.8	1726	11	Q8VC68	Q8VC68 mus muscucu
37	428.5	9.8	2622	11	Q70511	Q70511 rattus norv
38	427.5	9.7	416	10	Q94C42	Q94C42 triticum ae
39	427.5	9.7	1088	4	Q13484	Q13484 homo sapien
40	426	9.7	448	5	Q9V400	Q9V400 drosophila
41	426	9.7	961	5	Q86B19	Q86B19 dictyosteli
42	423.5	9.6	615	13	Q801N1	Q801N1 xenopus lae
43	423.5	9.6	1897	4	Q7Z3G4	Q7Z3G4 homo sapien
44	423.5	9.6	1762	11	Q89521	Q89521 rattus norv
45	420.5	9.6	1009	5	Q8SWY2	Q8SWY2 drosophila

ALIGNMENTS

RESULT 1

Q9Y2V6 PRELIMINARY; PRT; 835 AA.

AC Q9Y2V6;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein (Cardiac ankyrin repeat kinase).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RA Wei Y.J., Ding J.F., Xiong H., Zhou Y., Hui R.T., Liew C.C.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Jeyaseelan R.;
 RT "Cardiac Ankyrin Repeat Kinase (CARK)";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF116826; AAD29632.1; -;
 DR EMBL; AY303691; AAP72030.1; -;
 DR HSSP; Q00420; IAWC.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00023; ank; 10.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR01415; ANKYIN.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase.
 DR SMART; SM00248; ANK; 8.
 DR PROSITE; PS50088; ANK_REPEAT; 6.

DR PROSITE; PS00297; ANK REP REGION; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein; ANK repeat; ATP-binding; Kinase; Repeat;
KW Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 835 AA; 92850 MW; 3B21484B434F46B8 CRC64;

Query Match 100.0%; Score 4390; DB 4; Length 835;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGNYSRPTQCTDEWKKKVSSEYVITIERLEDDLOIKEXELTELNRNIFGSDFAFSKYNL 60
DB 1 MGNYSRPTQCTDEWKKKVSSEYVITIERLEDDLOIKEXELTELNRNIFGSDFAFSKYNL 60

QY 61 NYRTENGSLHLCCICCGGKSHIRTLMLKGLRSPRLTRNGFTALHLAVYKDAEILTSL 120
DB 61 NYRTENGSLHLCCICCGGKSHIRTLMLKGLRSPRLTRNGFTALHLAVYKDAEILTSL 120

QY 121 LHSAGADIQQVGYGGLTALHIAITAGHLEADVLLQHGANNVNIQDAVFTEPLHIAAYGHE 180
DB 121 LHSAGADIQQVGYGGLTALHIAITAGHLEADVLLQHGANNVNIQDAVFTEPLHIAAYGHE 180

QY 181 QVTRILLKFGADVNSGEVGRPLHLASAKGFNIAKLMEEGSKADVNAQDNEDHVPFLH 240
DB 181 QVTRILLKFGADVNSGEVGRPLHLASAKGFNIAKLMEEGSKADVNAQDNEDHVPFLH 240

QY 241 FCSRFGHHDIQVYLLQSDLEVQPHVNIYDGTPLHLACYNGKFEVAKETIIQISGTESLTK 300
DB 241 FCSRFGHHDIQVYLLQSDLEVQPHVNIYDGTPLHLACYNGKFEVAKETIIQISGTESLTK 300

QY 301 ENIFSETAFHSACTYKSIDLVKFLDQNVININHQGRDGHGTLHSAHYGHIRLVQFLL 360
DB 301 ENIFSETAFHSACTYKSIDLVKFLDQNVININHQGRDGHGTLHSAHYGHIRLVQFLL 360

QY 361 DNGADMLVACDPSSRSGEKDEQTCMLWAYEKGDHDAIVTLKHYKRPQDELPCNEYSQPG 420
DB 361 DNGADMLVACDPSSRSGEKDEQTCMLWAYEKGDHDAIVTLKHYKRPQDELPCNEYSQPG 420

QY 421 GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFHLQISEIFHEIIGSGSFGKVKYK 480
DB 421 GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFHLQISEIFHEIIGSGSFGKVKYK 480

QY 481 RCRNKIYAIKRYRANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLNDRPSQFAIVTQ 540
DB 481 RCRNKIYAIKRYRANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLNDRPSQFAIVTQ 540

QY 541 YISGSLFSLHLEQKRIIDLQSKLIIAVDVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG 600
DB 541 YISGSLFSLHLEQKRIIDLQSKLIIAVDVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG 600

QY 601 HAVVADFGESRFLQSLDEDNNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660
DB 601 HAVVADFGESRFLQSLDEDNNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660

QY 661 GEIPFAHLKPAADAAADYMHIRPPIGYSIPIKPSISLLIRGNACPEGRPEFESEVVMKLE 720
DB 661 GEIPFAHLKPAADAAADYMHIRPPIGYSIPIKPSISLLIRGNACPEGRPEFESEVVMKLE 720

QY 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780
DB 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780

QY 781 LSQSAGQYSSQGLSLEEMKRSLOVTPIDKYGVSDPMSSMHFHSNCRNSSFEDSS 835
DB 781 LSQSAGQYSSQGLSLEEMKRSLOVTPIDKYGVSDPMSSMHFHSNCRNSSFEDSS 835

RESULT 2

Q7TOP6

ID Q7TOP6

AC Q7TOP6;

DT 01-OCT-2003 (TRENBLrel. 25, Created)

PRELIMINARY;

PRT; 835 AA.

RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK052882; BAC35187.1; -;
 DR EMBL; AK084817; BAC39285.1; -;
 DR MGD; MGI:2443276; D830019J24Rik.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; ank; 3.
 DR SMART; SM00248; ANK; 3.
 DR PROSITE; PS0088; ANK_REPEAT; 1.
 DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 165 AA; 18372 MW; FA9C859545EB657D CRC64;
 Query Match 15.3%; Score 672; DB 11; Length 165;
 Best Local Similarity 85.3%; Pred. No. 2.4e-45;
 Matches 128; Conservative 7; Mismatches 15; Indels 0; Gaps 0;
 QY 1 MGNVSRPTQTCTDEWKKVSESYVITIERLEDDLOIKKELTELNRNIFGSDFAFSKVN 60
 DB 1 MGNVSRPTQTCTDEWKKVSESYVITIERLEDDLOIKKELTELNRNIFGSDFAFSKVN 60
 QY 61 NYRTENGLSLHLLCCICGGKSHIRTLMLKGLRPSRLTRNGFTALHLAVYKONAEITSL 120
 DB 61 NYRTENGLSLHLLCCACGGKSHIRPALMLKGLRPSRLTRNGFPALHLAVYKDSLEITSL 120
 QY 121 LHSAGDIOVQYGGHTALHIATIAHLEAA 150
 DB 121 LHSAGDVQVQYGGHTALHIAAGHPEVS 150
 RESULT 5
 ID Q22558 PRELIMINARY; PRT; 546 AA.
 AC Q22558;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN AT2G17700
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chang C.H., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RA "Full length cDNA of gene At2g17700 (GI:15227883).";
 RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RA "Arabidopsis Open Reading Frame (ORF) Clones";
 RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AY070086; AAL49781.1; -;

DR EMBL; AY096470; AAM20110.1; -;
 DR PIR; D84555; D84555.
 DR HSSP; P08631; IAD5.
 DR GO; GO:0016597; F:amino acid binding; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR002912; ACT.
 DR InterPro; IPR007019; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR01245; Tyr_pkinase.
 DR Pfam; PF01842; ACT; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; P00109; TYRKINASE.
 DR PRODOM; P000001; Prot_kinase; 1.
 DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Hypothetical protein; ATP-binding; Kinase;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 546 AA; 61509 MW; E74F30F60D91F426 CRC64;
 Query Match 11.1%; Score 486; DB 10; Length 546;
 Best Local Similarity 34.4%; Pred. No. 9.9e-30;
 Matches 114; Conservative 73; Mismatches 120; Indels 24; Gaps 8;
 QY 416 YSPGGGDSVSPSLGKIKSM--TKEKA-----DILLRAGLPS----HEHL 458
 DB 222 WSQETDGLRDALSKELKLDQPGSKQKSIFFEHDKSSNELIPACIEIPDTGDEWEI 281
 QY 459 QLSEIEPHEIIGSGSGKVKYGRNKKIVAIKRYRANTYCSKSDVDMFCREVSILCOLNH 518
 DB 282 DVTQTKIEKKVASSGYDLHRTYCSQEVAKFLKPD-R-VNNEMLREFSQEVFINRVRH 340
 QY 519 PCVIQFVAGCLNDPSQFAIVTQVIGSGSLFSLHLEOKRILDLQSKLIITAVYAKMEYLH 578
 DB 341 KXVQVQLGACTRSPT-LCIVTFMARGSIYDFLHKQKAFKLTLLKVALDVAKGMSYLH 399
 QY 579 NLTPILHRLNHNILLYEDGAVADFGESRFLQSLDEDNNTKPGNLRWVAPEVTO 638
 DB 400 Q--NNIHRDLKTANLMDHGLVKVADFGVARV--QIESGVMTAETGYRWVAPEVI-E 454
 QY 639 CTRYTIKADVFSALCWEILTEIPFAHLKPAAPAAADWAYHHIRPPIGYSIPKPISSLL 698
 DB 455 HKPNYHKADVFSVAIVLWELLTGDIPIYAFELTPIQAAVGVVQKGLRPPKIPKTHPKVKGLL 514
 QY 699 IRGNWACPEGRPFSEVVMKLECLCNIELM 729
 DB 515 ERCWHQDFEORPLFEIIEMLQOIMKEVNVV 545
 RESULT 6
 ID Q7XXN2 PRELIMINARY; PRT; 603 AA.
 AC Q7XXN2;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative serine/threonine-protein kinase ctrl.
 GN P0705811.7.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Oliveira A.C., Mattos L.T., Carvalho F.F., Shimano A., Zimmer P.D.,
 RA Malone G., Dellagostin G.;
 RT "Oryza sativa nipponbare genomic DNA, chromosome 9, PAC

```

RT clone:PO705E11. ";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006548; BAC79157.1; -.
KW Kinase.
SQ SEQUENCE 603 AA; 66329 MW; 78D406F3E2C3FB12 CRC64;

Query Match
Best Local Similarity 11.0%; Score 482; DB 10; Length 603;
Matches 130; Conservative 29.4%; Pred. No. 2.4e-29; Indels 40; Gaps 10;

QY 296 ESLTKENIFSTAHSACTYKSIDLVKFLDDQNVININHQGRD-----GHTGLHSA 347
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 348 CYHGHR-----LVQFLDNGADNLVACDPSSSGEKDEQCLMWAYEKGCHDAIVTL 401
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 210 TFTSTDKPKLLSELTSLGELGLNIQBAHAFSTNDGYSLDVFVVQGWDEETDLIESVR 269
QY 402 KHYKRPQDELPCNEYSQPGDGSVVSFSLGKIKMTKEKADILLRAGLPSHFHLQLS 461
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 270 KE-----IGKIDETQGWSTTHSSPVENMQIGENSAADHVEI PRDGASEWEIDVK 320
QY 462 EIEFHEITGSGSGFKVYKGRCKNIKVAIKRYANTYCSKSDVM----PCREVSIICQLNH 518
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 321 LILKFNKVASGSGDLVRGTYCSQDVAIKVLKP-----ERINADMQREFAQEVYIMRKVRH 376
QY 519 PCVIOFVGACLNDPSQFAIVTOYISGGSLFSLHLEOKRIILDQSLKLIIVADVAKGMEYLH 578
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 377 KNVVQFIAGCTKPPN-LCIVTEYNSGSGVDYLHKHGKFKPALLGVVMDYSKGSYLH 435
QY 579 NLTQPIIHRDLNSHNILLYEDGHAVADPGESFLOSLEDENNMTQFQNLRWMAVEVFTQ 638
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 436 Q--NIIHRDLKNTANLLMDENGTVKVAIDFGVARY--KAQSGVMVTAETGYRWMAPEVI-E 490
QY 639 CTRYTIKADVSYALCWEIITGPIPAHLKPAASAAADMAHYHRIPIGYSIPKPSISLL 698
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 491 HKPYDHKADVSFGILMMWELTGKPIEYLTPLQAAVGVVQKGRETIPKNAHAKLSSELL 550
QY 699 IGRGNACPEGRPEFSEVVVMKLE 720
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 551 QKCWQOEPAERPDFSEILETLQ 572

RESULT 7
Q9LNF8 PRELIMINARY; PRT; 445 AA.
ID Q9LNF8 AC
AC Q9LNF8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F16A14.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei Q., Chin C., Chioi J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwarz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F16A14 from chromosome
RT I.";
RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]

```

[illegible]

RESULT 8


```
Db 334 VDFGLSKIIKAHANDVYKMTGETSGSYRMAPEVFKH-RKYDKKVDVFSFAMILYEMLE 392
QY 661 GEIPFAHLPAADAAADMAVHHIRPPT---GYGIPKPISSLLIRGNACPEGRPFSEVVM 717
Db 393 GDPFFSNFPEYEAAYKVGSGH-RPFRSGFT--NELKELIELCWSGDHLRPSLEILK 449
QY 718 KLEE 721
Db 450 RLEK 453

RESULT 10
Q9X187 QX187 PRELIMINARY; PRT; 438 AA.
AC Q9X187
DT 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE F7A19.9 protein.
GN F7A19.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP STRAIN=cv. Columbia;
RC Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
RL EMBL; AC007576; AAD39286.1; -.
DR PIR; C86273; C86273.
DR HSP; Q00534; 1B18.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00023; ank; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00248; ANK; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ANK repeat; ATP-binding; Kinase; Repeat;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 438 AA; 49335 MW; C61F80724E24BBD2 CRC64;

Query Match 10.8%; Score 475; DB 10; Length 438;
Best Local Similarity 32.8%; Pred. No. 5.2e-29;
Matches 134; Conservative 73; Mismatches 155; Indels 46; Gaps 12;

QY 322 VKFLDQNVINHGQRDGTGLHSACVHGIRLVQFLDNGADMVLVACDPSRSSGEKD 381
Db 53 VRKLEEDPTLVHARDYDKRTPLHVASLHGWDVVKCLLEFGADV-----AQDRW 103
QY 362 EOTCLMAYEKGDAIVTLKHKKRPQDELPCNYSQPGDGSVSVPSPLGIKSMTKX 441
SQ SEQUENCE 438 AA; 49363 MW; 386FDA64886064A CRC64;
```


RA "Arabidopsis cDNA clones";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF462860; AAL58946.1; -
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 525 AA; 58841 MW; 88C2A6D6C6913856 CRC64;

Query Match 10.8%; Score 472; DB 10; Length 525;
Best Local Similarity 32.4%; Pred.No.1.2e-28;
Matches 122; Conservative 68; Mismatches 127; Indels 60; Gaps 13;

Qy 430 SPGLKIKSM-----TKEKADI-----LLLRAG-----LPSHFLQLSIEIFHEIIG 470
Db 155 SPSPGPIRLGLTIRQERVKSKDTGWSKLFNDTGRVSAVEASEFRVDMSKLFFGLKFA 214

Qy 471 SGSGFKYKRCNKIVAKRYA-----NTYKSDVD-MECREVSILCOLNHPVCVIOFV 525
Db 215 HGLYSRLHYKYEKAVAKLIITVDDONGCUGARLEKQFTKEVTLSSLRTHPNVIRFV 274

Qy 526 GACINDSQFAIVTQVIGSGSLSLHE-QKRILDLQSKLIIADVAKGMEYLNLTQPI 584
Db 275 GA-YKDPVPVYCVLTQVLPESGLASFLHKPENRSLPLKLIIEFVIDIARGVEYIH--SRRI 331

Qy 585 IHRDLNSHLLIYDGHVAVDFG----ERFLOSDEDNMTQPGNLRWMAPEVTQCTR 641
Db 332 IHRDLKPEVNLIDEEFLKIADFGIACEEYICDMLADD-----PGTYRWMAPEMKR-KP 385

Qy 642 YTIKADYFVALCLWETLTGEIIPFAHLKPAADAAADWYVHHIRPPTIGYSTPKISILLIRG 701
Db 386 HGRKADYFGLVLNENWAGAIYEDWNPIQAAFAVVKHNIKRIAPFGDCPVANKALIEQC 445

Qy 702 WNACPEGRPFSEVVMKLECLCNIELMSPASSNSGSLSPSSSDCL-----749
Db 446 WSVAPDKRPFQVTKVLEQFAISLE-----RKGNLNLSKSKICKDPRKGLKHWIQK 497

Qy 750 ---VNRGPGGRSHVAAL 763
Db 498 LQPVHAGGGGGSSSGL 514

RESULT 14
Q86WL6 PRELIMINARY; PRT; 570 AA.
ID Q86WL6
AC Q86WL6
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN AT4G35780
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AY03017; AAM13016.1; -
DR EMBL; AY128938; AAM91338.1; -
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF01842; ACT; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 570 AA; 64696 MW; D0961D218BE49344 CRC64;

RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Tripp M., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AY03017; AAM13016.1; -
DR EMBL; AY128938; AAM91338.1; -
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF01842; ACT; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 570 AA; 64696 MW; D0961D218BE49344 CRC64;

Query Match 10.7%; Score 469.5; DB 10; Length 570;
Best Local Similarity 32.2%; Pred.No.2.2e-28;
Matches 112; Conservative 74; Mismatches 120; Indels 42; Gaps 10;

Qy 388 WAVEKGDHDAIVTLKKHYKRPQDELPCNYSQPGDGSYVSPSLGKIKSMT-----REK 442
Db 228 WSQETETGLKDALKEIRFKDQ-PCS-----KQKSIFFPEHDKST 267

Qy 443 ADILLRAGLPS-----HFHLQSLSEIEFHEIISGSGFKYKRCNKIVAKRYANYTC 498
Db 268 NEILPACVEIPTDGTDEWIDMKQLKIEKKVACGSYGELEFRGTCSEVAILKLP---- 323

Qy 499 SKSDVDM---FCREVSILCOLNHPCVIOFVQVACLNDDPSQFAIVTQVIGSGSLFSLHEOK 555
Db 324 ERVNAEMLRFSQEVYIMKRVKHNVOFVIGACTRSPN-LCIVTEFTMGSIYDFLHKH 382

Qy 556 RILDQSKLIIADVAKGMEYLNLTQPIIHRDLNSHLLIYDGHVAVDFGSRFLQS 615
Db 383 GVFKIQSLKVALDVSQWNYLHQ--NNIHRDLKTANLLMDEHEVVKVADFQVARV--Q 438

Qy 616 LDEDNMTQPGNLRWMAPEVFTQCTRYTIKADYFVALCLWETLTGEIIPFAHLKPAADAA 675
Db 439 TEGSVMTAETGTYYRWMAPEVI-EHKPYDHRADVFAIVLWELLGELPYSYLTPLQAAV 497

Qy 676 DMAYHHRPPIGYSIPKISSLLIRGNACPEGRPFSEVVMKLECL 723
Db 498 GVQVQGLRPKIPKETHPKLTLEKQVQDPAIRPNFAEIIEMLNQLI 545

RESULT 15
Q86AT8 PRELIMINARY; PRT; 638 AA.
ID Q86AT8
AC Q86AT8
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to Dictyostelium discoideum (Slime mold). ankyrin repeat

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 2, 2004, 14:08:16 ; Search time 27 Seconds
(without alignments)
1610.318 Million cell updates/sec

Title: US-10-626-173-2
Perfect score: 4390
Sequence: 1 MGNYKSRPTQTCTDEWKKV.....PMSSMHFHSRNSSPEDSS 835

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	465	10.6	1584	1 KYK1_DICDI	P18160 dictyosteli
2	460.5	10.5	1880	1 ANK1_HUMAN	P16157 homo sapien
3	457.5	10.4	1862	1 ANK1_MOUSE	Q02357 mus musculu
4	450.5	10.3	410	1 P18161_DICDI	P18161 dictyosteli
5	441	10.0	1059	1 Y379_HUMAN	O15084 homo sapien
6	435	9.9	3924	1 ANK2_HUMAN	Q01484 homo sapien
7	427.5	9.7	4377	1 ANK3_HUMAN	Q12955 homo sapien
8	419	9.5	579	1 M3K7_MOUSE	Q62073 mus musculu
9	419	9.5	606	1 M3K7_HUMAN	Q43318 homo sapien
10	414	9.4	954	1 M3KA_HUMAN	Q02779 homo sapien
11	404.5	9.2	821	1 CTR1_ARATH	Q05609 arabidopsis
12	401	9.1	757	1 HT16_HYDAT	P53356 hydra atten
13	400.5	9.1	452	1 ILK1_HUMAN	Q13418 homo sapien
14	396	9.0	452	1 ILK_MOUSE	O55222 mus musculu
15	390.5	8.9	452	1 ILK3_HUMAN	P57043 homo sapien
16	388	8.8	451	1 ILK_CAVPO	P57044 cavia porce
17	375	8.5	832	1 ANR3_HUMAN	P57078 homo sapien
18	367.5	8.4	625	1 ITK_MOUSE	Q03526 mus musculu
19	364	8.3	620	1 ITK_HUMAN	Q08881 homo sapien
20	358.5	8.2	467	1 MATK_RAT	P41243 rattus norv
21	357.5	8.1	822	1 FER_HUMAN	P16591 homo sapien
22	356	8.1	859	1 M3KC_HUMAN	Q12852 homo sapien
23	356	8.1	888	1 M3KC_MOUSE	Q60700 mus musculu
24	355	8.1	394	1 M3K9_HUMAN	P80192 homo sapien
25	354	8.1	768	1 YB23_HUMAN	Q09137 homo sapien
26	354	8.1	888	1 M3KC_RAT	Q63796 rattus norv
27	353	8.0	509	1 STK_HYDAT	P17713 hydra atten
28	353	8.0	735	1 RNSA_MOUSE	Q05921 mus musculu
29	352.5	8.0	1520	1 ABL_DROME	P00522 drosophila
30	351	8.0	647	1 RAF1_CHICK	P05625 gallus gall
31	349.5	8.0	505	1 MATK_MOUSE	P41242 mus musculu
32	349	7.9	450	1 CSK_CHICK	P41239 gallus gall
33	348.5	7.9	638	1 RAF1_XENLA	P09560 xenopus lae

34	346.5	7.9	552	1 SR64_DROME	P00528 drosophila
35	345.5	7.9	648	1 RAF1_RAT	P11345 rattus norv
36	345	7.9	505	1 SRK1_SPOLA	P42686 sporgilla 1
37	344	7.8	380	1 MIL_AVIMH	P00531 avian retro
38	342.5	7.8	648	1 RAF1_MOUSE	Q09n57 mus musculu
39	342	7.8	539	1 RIK2_MOUSE	P58801 mus musculu
40	341	7.8	1442	1 DAK1_MOUSE	Q80ye7 mus musculu
41	340	7.7	1122	1 TIE2_MOUSE	Q02858 mus musculu
42	340	7.7	1124	1 TIE2_HUMAN	Q02763 homo sapien
43	340	7.7	1125	1 TIE2_BOVIN	Q06807 bos taurus
44	339	7.7	1432	1 DAK1_HUMAN	P53355 homo sapien
45	337	7.7	527	1 TXK_HUMAN	P42681 homo sapien

ALIGNMENTS

RESULT 1

KYK1_DICDI
ID KYK1_DICDI STANDARD; PRT; 1584 AA.
AC P18160:
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Non-receptor tyrosine kinase spore lysis A (SC 2.7.1.112) (Tyrosine-
protein kinase 1).
DE SPLA OR PYKA OR DPYK1.
GN Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JH10;
RX MEDLINE=970531827; PubMed=8898241;
RA Nuckolls G.H., Oshervov N., Loomis W.F., Spudich J.A.;
RT "the Dictyostelium dual-specificity kinase spla is essential for
spore differentiation.";
RL Development 122:3295-3305(1996).
RN [2]
RP SEQUENCE OF 1248-1584 FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
Dictyostelium discoideum.";
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -!- FUNCTION: Essential for spore differentiation.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -!- DEVELOPMENTAL STAGE: Expressed throughout development with a peak
during the mound stage of morphogenesis.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases.
CC -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U32174; AB41125.1; --
EMBL; M33785; AAA3202.1; --
PIR; T18276; T18276.
DictyBase; DDB0001716; spla.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR001660; SAM.
InterPro; IPR003877; SPRY_receptor.
InterPro; IPR001245; Tyr_kinase.
InterPro; IPR008266; Tyr_kinase_AS.
Pfam; PF00069; kinase; 1.
Pfam; PF00536; SAM; 1.
Pfam; PF00622; SPRY; 3.

```

DR PRINTS; PRO0109; TYRKINASE.
DR ProDom; PD000001; Prot_Kinase; 1.
DR SMART; SM00454; SAM; 1_Kinase; 1.
DR SMART; SM00449; SPRY; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS0105; SAM_DOMAIN; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT DOMAIN 908 972 SAM.
FT DOMAIN 403 420 POLY-ASN.
FT DOMAIN 428 435 POLY-THR.
FT DOMAIN 449 480 POLY-ASN.
FT DOMAIN 483 491 POLY-ASN.
FT DOMAIN 494 508 POLY-ASN.
FT DOMAIN 512 532 POLY-ASN.
FT DOMAIN 536 600 POLY-ASN.
FT DOMAIN 808 811 POLY-PHE.
FT DOMAIN 1026 1029 POLY-SER.
FT DOMAIN 1195 1210 POLY-ASN.
FT DOMAIN 1215 1220 POLY-GLN.
FT DOMAIN 1224 1233 POLY-GLN.
FT DOMAIN 1266 1274 POLY-PRO.
FT DOMAIN 1289 1561 PROTEIN_KINASE.
FT NP_BIND 1295 1303 ATP (BY SIMILARITY).
FT BINDING 1316 1316 ATP (BY SIMILARITY).
FT ACT_SITE 1417 1417 BY SIMILARITY.
FT CONFLICT 1248 1248 D -> R (IN REF. 2).
FT CONFLICT 1435 1435 V -> L (IN REF. 2).
SQ SEQUENCE 1584 AA; 174304 MW; 5D1589458D8E01E3 CRC64;

Query Match 10.6%; Score 465; DB 1; Length 1584;
Best Local Similarity 37.7%; Pred. No. 1.4e-20;
Matches 109; Conservative 52; Mismatches 100; Indels 28; Gaps 6;

QY 454 SHFLQLGEIFHFHIGSGFGKVKGRCKRNIKVAIKRYRANTYCSKSDVDMFCREVSIL 513
DB 1280 SEYDFNELEFGQTIGKGFGEVKGRYRETDVAIKIYRDQPKTKSLVMQNEVGIL 1339

QY 514 COLNHPVQIFVGAC-LNDPSQFAIVQYISGSL-----PSLLHEQKRLDLOSKLI 565
DB 1340 SKLRHPNVVQFLGACTAGGEDHHCHIVTEWMGGSLRQFLTDHFNLLQNPHI-----RLK 1394

QY 566 IAVDVAKGMEVHLNLTQPIIHRDLNSHNILL-----YEDGHAVVADFGESRF 612
DB 1395 LALDIKAGNVLHGHTWTPPIILHRDLSSRNILLDHPNPVVSRRQDKICKIDFGLSR- 1453

QY 613 LOSLDDNNTKQPNLRMAPEVFTQCTRYTIKADVPSYALCWLWEILTGIPFAHLKPAA 672
DB 1454 LKKEQASQNTQSVGCIPTMAPEVFGDSN-SEKSDVTSYGMVLFELLITSDPEQDDMKPKR 1512

QY 673 AAADWAYHHIPPGYSPKPISSILLRGWNAACEGPEPESEVVMKLEE 721
DB 1513 MAHLAAYSAPPIPLITTSKWKELITQCWDSNPSRPTFKQIVHLKE 1561

RESULT 2
ANK1_HUMAN
ID ANK1_HUMAN STANDARD; PRT; 1880 AA.
AC P16157;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ankyrin 1 (Erythrocyte ankyrin) (Ankyrin R).
GN ANK1 OR ANK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.
RC TISSUE=Hematopoietic;
RX MEDLINE=90158630; PubMed=2137557;

```

Lux S.E., John K.M., Bennett V.;
 "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated
 structure with homology to tissue-differentiation and cell-cycle
 control proteins";
 Nature 344:36-42 (1990).
 [2]
 SEQUENCE FROM N.A.
 RX MEDLINE=90175370; PubMed=1689849;
 RA Lambert S., Yu H., Prchal J.T., Lawler J., Ruff P., Speicher D.,
 Cheung M.C., Kan Y.W., Palek J.;
 "cDNA sequence for human erythrocyte ankyrin";
 Proc. Natl. Acad. Sci. U.S.A. 87:1730-1734 (1990).
 [3]
 VARIANT HS ILE-462.
 RX MEDLINE=96225450; PubMed=8640229;
 RA Eber S.W., Gonzalez J.M., Lux M.L., Scarpa A.L., Tse W.T.,
 Dornwell M., Herbers J., Kugler W., Oerzcan R., Pekrun A.,
 Gallagher P.G., Schroeter W., Forget B.G., Lux S.E.;
 "Ankyrin-1 mutations are a major cause of dominant and recessive
 hereditary spherocytosis";
 Nat. Genet. 13:214-218 (1996).
 CC -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
 elements; bind to the erythrocyte membrane protein band 4.2, to
 Na-K ATPase to the lymphocyte membrane protein Gp95, and to the
 cytoskeletal proteins fodrin, tubulin, vimentin and desmin.
 CC Erythrocyte ankyrins also link spectrin (beta chain) to the
 cytoplasmic domain of the erythrocytes anion exchange protein;
 they retain most or all of these binding functions.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC PLASMA
 MEMBRANE.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1; Synonyms=2.1;
 CC IsoId=P16157-1; Sequence=Displayed;
 CC Name=2; Synonyms=2.2;
 CC IsoId=P16157-2; Sequence=VSP_000264; VSP_000265;
 CC Name=3;
 CC IsoId=P16157-3; Sequence=VSP_000266;
 CC -!- PTM: Regulated by phosphorylation.
 CC -!- PTM: Palmitoylated.
 CC -!- DISEASE: Defects in ANK1 are a cause of hereditary spherocytosis
 (HS) [MIM:182900]. Inheritance can be autosomal dominant or
 recessive.
 CC -!- SIMILARITY: Contains 23 ANK repeats.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 EMBL; X16609; CAA34610.1; -.
 DR EMBL; M28880; AAA51732.1; -.
 DR PIR; A35049; A35049.
 DR PIR; S08275; SJHUK.
 DR HSSP; Q00420; LAWC.
 DR Genew; HGNC:492; ANK1.
 DR MIM; 182900; -.
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR000488; Death.
 DR InterPro; IPR000906; ZUS.
 DR Pfam; PF00023; ank; 24.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00791; ZUS; 1.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 21.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00218; ZUS; 1.

CC -!- PTM: Regulated by phosphorylation (By similarity).
 CC -!- PTM: Acylated by palmitic acid group(s) (By similarity).
 CC -!- SIMILARITY: Contains 23 ANK repeats.
 CC -!- SIMILARITY: Contains 1 death domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M84756; AAA37236.1; .
 DR PIR; I49502; I49502.
 DR HSSP; Q00420; 1ANC.
 DR MGD; MGI:88024; Ank1.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR000488; Death.
 DR Pfam; PF00023; ank; 24.
 DR Pfam; PF00531; death; 1.
 DR Pfam; PF00791; ZUS; 1.
 DR PRINTS; PRO14145; ANKYRIN.
 DR SMART; SM00248; ANK; 22.
 DR SMART; SM00005; DEATH; 1.
 DR SMART; SM00218; ZUS; 1.
 DR PROSITE; PS50088; ANK REPEAT; 20.
 DR PROSITE; PS50297; ANK REP REGION; 1.
 DR PROSITE; PS50017; DEATH DOMAIN; 1.
 KW Cytokeleton; Repeat; ANK repeat; Phosphorylation; Lipoprotein.
 FT DOMAIN 1 827
 FT 828 1386
 FT DOMAIN 1387 1862
 FT 40 69
 FT REPEAT 73 102
 FT REPEAT 106 135
 FT REPEAT 139 168
 FT REPEAT 170 197
 FT REPEAT 201 230
 FT REPEAT 234 263
 FT REPEAT 267 296
 FT REPEAT 300 329
 FT REPEAT 333 362
 FT REPEAT 366 395
 FT REPEAT 399 428
 FT REPEAT 432 461
 FT REPEAT 465 494
 FT REPEAT 498 527
 FT REPEAT 531 560
 FT REPEAT 564 593
 FT REPEAT 597 626
 FT REPEAT 630 659
 FT REPEAT 663 692
 FT REPEAT 696 725
 FT REPEAT 729 758
 FT REPEAT 762 791
 FT DOMAIN 1399 1483
 FT DEATH.
 SQ SEQUENCE 1862 AA; 204242 MW; AE6B5B5B29001E5 CRC64;

Query Match 10.4%; Score 457.5; DB 1; Length 1862;
 Best Local Similarity 28.7%; Pred. No. 4.9e-20;
 Matches 151; Conservative 66; Mismatches 171; Indels 139; Gaps 17;
 QY 60 LNYRTE-----NGLSLHLCCICGGKKSHRTLMKGLRPSRLTRNGFTALFLAVYKDN 113
 DB 321 LQYNAEIDDTLDHLTLPLHVAHCGHHRV-AKVLLDKGAKPNSRALNGFTPLHIACKGNH 379
 QY 114 AELTSLHSGADIQQGVYGGTLTHIATAGHLEAADVLQHGANNVTDQAVFFTLPLHI 173

DB 360 IRVMELLKTKGASDAVTESGTLPLHVASFMGHLPIVKNLQRGASPNVNVKVEITPLHM 439
 QY 174 AAYGHEQVTRLLKFGADVNVSGVDRLPHLASAKGFLNTAKILMEGSGKADVNAQDN 233
 DB 440 AARAGHTEVAKYLLQNKAKANAKAKODQTEPLHCAARIGHTGMVKLLLENG--ASPNLAT 497
 QY 234 EDHVPFLFCRFGH-----HDIVKY-----LLQSD-- 258
 DB 498 AGHTPLHTAREGHVDLTALLSKEASQACMTKGGTFLHVAKYKVELAELLEHDHDAH 557
 QY 259 -----LEV-----OPHVNIYGDTPHLACYNCKFEVAKEI 289
 DB 558 PNAAGKNGLTPLHVAHNNLDIVKLLPRGGSPHSPAWNGYTPLHIAAKQNKIEVARSL 617
 QY 290 IQISGTESLTKENIFSETAFHSACTYCKSIDLVKFLLDQNV----- 330
 DB 618 LQYCG--SANAESVQGVTPHLAAQSGHT-EMWALLSKQANGNLGNKSGTLPLHLVSGE 674
 QY 331 -----ININHQGRDGTGLHSACYHGHIRLVQFLIDNGADNMLVACDPSSRS 377
 DB 675 GHVLVADVLIKHGVTVDATTRMGYTPHLVASHYGNIKLVKFLQLHQADV-----A 725
 QY 378 GEKDEQTCIMWAYEKGHDAIVT-LLKHXYRQDELPCNEYSPQGGD-----SYVSV 429
 DB 726 KTLGYSPLHQAQGGHTDIVTLLKNGASP-----NEVSSNGTTPPLAIAKRLGYISVT 779
 QY 430 SPLGKIKSMTKERADIILL---RAGLPSPHF--LQLSEIEFHEIIGS 471
 DB 780 DVL---KVVTDETSVLVSKHMSYPTVETDEILDVSEDEGDELVS 823
 RESULT 4
 ID KYK2 DICI STANDARD; PRT; 410 AA.
 AC P18161;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Tyrosine-protein kinase 2 (EC 2.7.1.112) (Fragment).
 GN SPLB OR PYXB OR DPKY2.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90287147; PubMed=1972546;
 RX Tan J.L., Spudich J.A.;
 RT "Developmentally regulated protein-tyrosine kinase genes in
 Dictyostelium discoideum";
 RL Mol. Cell. Biol. 10:3578-3583 (1990).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 tyrosine phosphate.
 CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M33784; AAA33203.1; .
 DR PIR; B35670; B35670.
 DR HSSP; P08631; IADS.
 DR DictyBase; DDB0002037; splB.
 DR InterPro; IPR000719; Tyr_kinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TIR; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT NON_TER 1
FT DOMAIN 108 381 PROTEIN KINASE.
FT NP_BIND 114 122 ATP (BY SIMILARITY).
FT BINDING 135 135 ATP (BY SIMILARITY).
FT ACT_SITE 232 232 BY SIMILARITY.
SQ SEQUENCE 410 AA; 46386 MW; E93918B6059AEC1 CRC64;

Query Match 10.3%; Score 450.5; DB 1; Length 410;
Best Local Similarity 31.1%; Pred. No. 1.8e-20;
Matches 131; Conservative 80; Mismatches 145; Indels 65; Gaps 14;

QY 365 DMNIVACDPSRSGSDEQCLMWAYEKGDHAIIVT---LLKHVYRQDELPCNEYSPQGG 421
DB 11 DITLVCDNPDSTKEKNVS-----NTSSIIISASLNLRH-----ITPNSHMRPG 55

QY 422 DG-SYVSPGLGK-----IKMTKEKADILLRAGL-PSHFLQLSIEBHEIIG 470
DB 56 RSISESLIMSPINKESLNDIQRATIESEKIKTKPEELKSLIGREYIIINDIQTQKVG 115

QY 471 SGSPGKVKGRCKNKAIVAKRYRANTYCSKSDVDMFCREVSILCOLNHPCVQFVGACLN 530
DB 116 EGAFSEVWGWGKIHWAIKKIKI-IGDEQFERFIREVQNLKGNHONVWFIGACYK 174

QY 531 DPSQFAIVTYISGGSFSLHE-----QKRIILDOSKLIIVDVAKGMEYLNLTQPII 585
DB 175 -PA-CIITEYAGGSLYNILHNPNSSTPKVYSFPLVLMKQATDMALGLHLHSIT--IV 229

QY 586 HRDLNSHILLYEDHAVADFGESRFLQSLDEDNMTKQPG---NLRWAPVEVFTQCTRY 642
DB 230 HRDLTSQNLIDELGNIKISDFGLS--AEKRSQMTWTGGTCNPRWPPPELTKNLGHY 287

QY 643 TIKADVSFALCLWEILTGEIPFAHLKPAADMAHYHRRPPIGYSIKPKISSLLIRGW 702
DB 288 SEKVDVYCFSLVWEILITGIPSLDGSQBSQAVAGLRPIPEYCDPELKLITQCM 347

QY 703 NACPEGRPEFSEVYKLEECUN-----IELMSPASNSGSLSP 742
DB 348 EADPNDRPPFTYIVNKLKEISWNNPIGFVSDQFYQYSEPTPTPLALSNSGSSISLSP 407

QY 743 S 743
DB 408 T 408

RESULT 5
Y379 HUMAN
ID Y379 HUMAN STANDARD; PRT; 1059 AA.
AC O15084;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical ANK-repeat protein KIAA0379 (Fragment).
GN KIAA0379.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain.
RX MEDLINE=9734984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
CC -!- SIMILARITY: Contains at least 27 ANK repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).

EMBL; AB002377; BAA20833.2; -.
HSP; P80144; 2MYO.
InterPro: IPR002110; ANK.
Pfam: PF00023; ank; 28.
PRINTS; PR01415; ANKYRIN.
SMART; SM00248; ANK; 28.
DR PROSITE; PS50088; ANK_REPEAT; 24.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Hypothetical protein; Repeat; ANK repeat.
FT NON_TER 1 75 ANK 1.
FT REPEAT 46 75 ANK 2.
FT REPEAT 79 108 ANK 3.
FT REPEAT 112 141 ANK 4.
FT REPEAT 145 174 ANK 5.
FT REPEAT 178 207 ANK 6.
FT REPEAT 211 240 ANK 7.
FT REPEAT 244 273 ANK 8.
FT REPEAT 277 307 ANK 9.
FT REPEAT 311 340 ANK 10.
FT REPEAT 344 373 ANK 11.
FT REPEAT 377 406 ANK 12.
FT REPEAT 410 439 ANK 13.
FT REPEAT 443 472 ANK 14.
FT REPEAT 476 506 ANK 15.
FT REPEAT 510 540 ANK 16.
FT REPEAT 555 584 ANK 17.
FT REPEAT 588 617 ANK 18.
FT REPEAT 622 651 ANK 19.
FT REPEAT 658 687 ANK 20.
FT REPEAT 691 720 ANK 21.
FT REPEAT 724 753 ANK 22.
FT REPEAT 761 790 ANK 23.
FT REPEAT 793 823 ANK 24.
FT REPEAT 828 857 ANK 25.
FT REPEAT 861 891 ANK 26.
FT REPEAT 895 924 ANK 27.
FT REPEAT 931 960 ANK 28.
SQ SEQUENCE 1059 AA; 113465 MW; C1F55E6CFE494770 CRC64;

Query Match 10.0%; Score 441; DB 1; Length 1059;
Best Local Similarity 24.2%; Pred. No. 2.4e-19;
Matches 214; Conservative 120; Mismatches 360; Indels 190; Gaps 40;

QY 21 SESVITIERLEDDLOIKEKELTELNRNIFGSD-----AFSKVNLNRYRTENGSLHLHC 74
DB 2 AEATAMAFKLIRQPSL-----VQAFNGDPDEVRALIFKEDVNFQDNEKRTPLHAA 54
QY 75 CIOGKSHIRITMLKGLRPSRLTRNGFTALHVAIVYKDNAELITSLHSGADIQQVGYGG 134
DB 55 AYLGDABE-ITELLISGARVNAKSKWLTPLHRAVASCSEEAQVLLKHSADVNARDKNW 113
QY 135 LTALHIAITAGHLEADVLIQHGANNVNIQDAVFTPLHTAAVYGHQVTRLLIKFGADV 194
DB 114 QTPHIAAANKACAEALPVLSSNVNVDRAQRTALHRAAFSGHGMVKLLLSRGANIN 173
QY 195 VSGEVGDRPLHLASAKGLNIAKLLMEESKADVNADQNEHDVPLHFCRFGHHDIKVYL 254
DB 174 AFDKDRRAIHWAAYMGHIEVVKLLVSHG--AEVCKDKKSYTFLHAAASSGMISVVKYL 231
QY 255 LOSLEVQPHVNNIYGTPLHLACYNGKFEVAKETIQISGTSLTKEITFSETAFHSACT 314
DB 232 L--DLGVDMEPNAYGNTPHLVACYNGQDVVNNELIDCGAIVNQKNEKGTPLHFAAST 289
QY 315 YGKSIDLKVFLLDQNVININHGCRDGTGLHSACYGHIRLVQFLDNGADMNLVACDPS 374

```
Db 290 HGA---LCLLELVGNGADVNMKSKDKTPLHMTALHGRFSRSQIIOSGA---VIDCE-- 341
Qy 375 RSSEKDEQTCMLMAYEKGDHDAIVTLKHXYKRPDELPCEYSPQGGD---GSYVSVPS 430
Db 342 ----DKNGNTPLHIAARYGHELLI-----NTLITSGADTAKRGHIGMFPL 382
Qy 431 PLGKIKSMYKXADILLRAGLPSHFLQLSEIEP-----HEIIGSGFGYKVGCRNK 485
Db 383 HLAALSGFSD-----CCNKLLSSGFDIDTPD-DGRTCLHAAAGNGL-----ECNLN 429
Qy 486 IVAIKRYRANTYCSKSDVDMFCRE---VSILCOLNHPCVQFV---GACILND-----P 532
Db 430 LL-----NTGADFNNKDKFGRSPLHYAAANC--NYOCLPALVSGSASVNDLDERGCTP 480
Qy 533 SOPAIVTOYISGSLFSLHBEQR--ILDQSLKLIIVDVAKMGMEYHNLTOPIIHRDLN 590
Db 481 LHVA-ATSDTDGKCLEYLRNDANFCIRDKQYNNAVHYSAAYG---HRLCQLIASE-T 534
Qy 591 SHNIIYEDGHAVADFGESRFLQSLDEDMTKPGNL-----RWMAPVFTQCTRYTIK 645
Db 535 PLDLVNETSGTDLSD-----SDNRATISPLHLAAVGHGHQALEVLVQS---LLD 581
Qy 646 ADVFSVALCLWILGEIP--FAHLKPAAAAADMAVHHIRPPIGYI-----PKPIS 695
Db 582 LDVRN-----SSGORTPLDLAFAFGHVECDVLNQ-----GASILVKDYILKRTPIH 628
Qy 696 SILLIRGNACPEGRPEFSEFVVMKLECLCNIELMSPASSNSSGSLSP-----SSSDC- 748
Db 629 AATNGHSEC-----LRLIGNAEPQNAVDIQNGQTPMLSLVNLGHTDCV 675
Qy 749 --LVNRGSP-----GRS--HVAALRSRFEVAL-----NARSVAALSQSA 785
Db 676 YSLNKGANVDADKWKGRALTALRGAVTGHEECVDALLQHGAKCLLRDSRGRTPIHLSAAC 735
Qy 786 QYSSOGLSLEEMKESLQYTP--IDKYGVSDPMSSMHFHSRNL 827
Db 736 GHIGVLG-ALLOSAASMDANPATADNHGY-----TALHW-ACYN 772

RESULT 6
ANK2 HUMAN
ID ANK2 HUMAN STANDARD; PRT; 3924 AA.
AC Q01484; C01485;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ankyrin 2 (Brain ankyrin) (Ankyrin B) (Ankyrin, nonerythroid).
GN ANK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3).
RC TISSUE=Brain stem;
RX MEDLINE=91102466; PubMed=1830053;
RA Otto E., Kunimoto M., McLaughlin T., Bennett V.;
RT "Isolation and characterization of cDNAs encoding human brain
RN ankyrins reveal a family of alternatively spliced genes.";
RN J. Cell Biol. 114:241-253(1991).
RP [2]
RP REVISIONS.
RA Carpenter S.;
RN Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain stem;
RX MEDLINE=94075409; PubMed=8253844;
RA Chan W., Kordeli B., Bennett V.;
RT "440-kD ankyrinB: structure of the major developmentally regulated
RN domain and selective localization in unmyelinated axons.";
RN J. Cell Biol. 123:1463-1473(1993).
RN [4]
```

```
RP SEQUENCE OF 463-495 FROM N.A.
MEDLINE=92009921; PubMed=1833308;
RX Tse W.T., Menninger J.C., Yang-Feng T.L., Francke U., Sahr K.E.,
Lux S.E., Ward D.C., Forget B.G.;
RT "Isolation and chromosomal localization of a novel nonerythroid
ankyrin gene.";
RL Genomics 10:858-866(1991).
CC -!- FUNCTION: Attach integral membrane proteins to cytoskeletal
elements. Also bind to cytoskeletal proteins.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q01484-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q01484-2; Sequence=VSP_000267; VSP_000268;
CC Name=3;
CC IsoId=Q01484-3; Sequence=VSP_000268;
CC -!- TISSUE SPECIFICITY: Plasma membrane of neurons as well as glial
cells throughout the brain.
CC -!- PTM: Phosphorylated at multiple sites by different protein kinases
and each phosphorylation event regulates the protein's structure
and function (Potential).
CC -!- SIMILARITY: Contains 23 ANK repeats.
CC -!- SIMILARITY: Contains 1 death domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
EMBL; X56957; CAA40278.1; -
EMBL; X56958; CAA40279.2; -
EMBL; Z26634; CAB42644.1; -
EMBL; M37123; AAA62828.1; -
FIR; S37431; S37431.
HSSP; P42771; 1DC2.
Gene; HGNC:493; ANK2.
MIM; 106410; -
InterPro; IPR002110; ANK.
InterPro; IPR000488; Death.
InterPro; IPR000906; ZU5.
Pfam; PF00023; ank; 24.
Pfam; PF00531; death; 1.
Pfam; PF00791; ZU5; 1.
PRINTS; PR01415; ANKYRIN.
SMART; SM00248; ANK; 22.
SMART; SM00005; DEATH; 1.
SMART; SM00218; ZU5; 1.
PROSITE; PS50086; ANK REPEAT; 20.
PROSITE; PS50297; ANK REP REGION; 1.
PROSITE; PS50017; DEATH DOMAIN; 1.
Cytoskeleton; Alternative splicing; Repeat; ANK repeat;
KW Phosphorylation.
FT REPEAT 63 92 ANK 1.
FT REPEAT 96 125 ANK 2.
FT REPEAT 129 158 ANK 3.
FT REPEAT 162 191 ANK 4.
FT REPEAT 193 220 ANK 5.
FT REPEAT 232 261 ANK 6.
FT REPEAT 265 294 ANK 7.
FT REPEAT 298 327 ANK 8.
FT REPEAT 331 360 ANK 9.
FT REPEAT 364 393 ANK 10.
FT REPEAT 397 426 ANK 11.
FT REPEAT 430 459 ANK 12.
FT REPEAT 463 492 ANK 13.
FT REPEAT 496 525 ANK 14.
FT REPEAT 529 558 ANK 15.
FT REPEAT 591 591 ANK 16.
FT REPEAT 595 624 ANK 17.
```

REPEAT	628	657	ANK 18.
FT	REPEAT	661	ANK 19.
FT	REPEAT	694	ANK 20.
FT	REPEAT	727	ANK 21.
FT	REPEAT	760	ANK 22.
FT	REPEAT	793	ANK 23.
FT	DOMAIN	1773	REPEAT-RICH REGION.
FT	REPEAT	1773	REPEAT A.
FT	REPEAT	1785	REPEAT A.
FT	REPEAT	1797	REPEAT A.
FT	REPEAT	1809	REPEAT A.
FT	REPEAT	1821	REPEAT A.
FT	REPEAT	1833	REPEAT A.
FT	REPEAT	1845	REPEAT A.
FT	REPEAT	1857	REPEAT A.
FT	REPEAT	1868	REPEAT A.
FT	REPEAT	1880	REPEAT A.
FT	REPEAT	1891	REPEAT A.
FT	REPEAT	1902	REPEAT A. (APPROXIMATE).
FT	REPEAT	1903	REPEAT A. (APPROXIMATE).
FT	REPEAT	1914	REPEAT A.
FT	REPEAT	1915	REPEAT A.
FT	REPEAT	1926	REPEAT A.
FT	REPEAT	1927	REPEAT A.
FT	REPEAT	1938	REPEAT A.
FT	REPEAT	1950	REPEAT A.
FT	DOMAIN	3536	DEATH.
FT	VARSPPLIC	3536	Q -> QFLGKHLPTAPPPLNEGSLVSRILQLGPPGTK (in isoform 2).
FT	VARSPPLIC	1039	/FTID=VSP 000267.
FT	VARSPPLIC	1444	Missing (in isoform 2 and isoform 3).
FT	CONFLICT	475	GO -> PE (IN REF. 4).
FT	CONFLICT	971	I -> S (IN REF. 1).
FT	CONFLICT	3581	QY -> HA (IN REF. 1).
FT	CONFLICT	3586	I -> Y (IN REF. 1).
FT	SEQUENCE	3924	AA; 430337 MW; 52AC496C428E29D2 CRC64;

Query Match

Best Local Similarity

Matches 133; Conservative

9.9%; Score 435; DB 1; Length 3924;

27.9%; Pred. No. 3e-18;

71; Mismatches 187; Indels 86; Gaps 12;

Qy	60	LNVTENGSLHLCCICGCKSHIRTLMLKGLRPSRLTRNGFTALHLAVYKDNALITS	119
Db	292	IDAKTRDGLTFLH-CAARSHDQVVELLERGAPILARTKNGLSPLHMAAQDHVECVKH	350
Qy	120	LLHSGADIQGVGGTLTALHIATTAHLEAADVLLQHGANNVIOAVFPTPLHIAAYGH	179
Db	351	LLQHKAPVDDVTLDYLTALHVAACHGHYRVTKLLDKRANPNARALNGFPLHIAACKQR	410
Qy	180	EOVTRLLKFGADVNVSGEVDRLPHLASAKGFLNIAKLIMEEGSKAD	227
Db	411	IKWELLVKYGASIQATESGLTPIHVAAPFMGHLNIVLLQLNGASPDVTNIRGETALHM	470
Qy	228	-----VNAQDNEDHVLPHFCRSRFGHHDIVKYLLOSLDLEVPQHVNI	268
Db	471	AARAGQEVVVRCLLRNGALVDASAREEOTPLHTASRLGKTEIVQLLQH--WAHPDAATT	528
Qy	269	YGDPTPLHACYNGKFEVAKETIIOISGTESLTENIFSETAFHSACTYKGSIDLVRFLLDQ	328
Db	529	NGYTPPLHISAREQVDVASVLLRAGAAHSLATKGF--TFLHVAAYG-SLDVAKLLQOR	585
Qy	329	NVINNHQGRDGTGLHSACYHGHIRLVQFLDNGADMLNACD-----PSSRSGEKDEQ--	383
Db	586	RAA-ADSAGKNGLTPLHVAAYHYNQKVALLEKSGASPHATKNGYTPPLHIAAKGNQMI	644
Qy	384	-----TCLMWAYEKGHDAIVTLLKHLKHYRQDELPCNVEYSQPGGDS	424
Db	645	ASTLLNLYGAETNIVTKGVTPPLHLASQEGHTDMVTL-----LKGANI	688
Qy	425	YVSVPSPLGKIKMTKEK----ADILLRAGLPSHFHLQSEIEFHEIIGSSSGFKV	477
Db	689	HMSTKSGLTSLHAAQEDKYNVADI-LTKHGADQDAHTKLG---YTPFLIVACHYGNV	741

RESULT 7

ANK3 HUMAN

RESULT 7
ANK3_HUMAN

ID	ANK3 HUMAN	STANDARD;	PRT; 4377 AA.
AC	Q12955;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	15-MAR-2004 (Rel. 43, Last annotation update)		
DE	Ankyrin 3 (ANK-3) (Ankyrin G).		
GN	ANK3.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RP	SEQUENCE FROM N.A.		
RN	TISSUE=Brain stem;		
RC	MEDLINE=95138209; PubMed=7836469;		
RX	Kordeli E., Lambert S., Bennett V.,		
RA	"Ankyrin. A new ankyrin gene with neural-specific isoforms localized		
RT	at the axonal initial segment and node of Ranvier.";		
RL	J. Biol. Chem. 270:2352-2359(1995).		
CC	-I- FUNCTION: Membrane-cytoskeleton linker. The neural-specific		
CC	isoforms may participate in the maintenance/targeting of ion		
CC	channels and cell adhesion molecules at the nodes of Ranvier and		
CC	axonal initial segments.		
CC	-I- SUBUNIT: Neural-specific isoforms may be a constituent of a		
CC	neurofascin/NRCAM/ankyrin G complex.		
CC	-I- ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=1;		
CC	Comment=A number of isoforms are produced;		
CC	Name=1; Synonyms=480-kDa isoform;		
CC	Isoid=Q12955-1; Sequence=Displayed;		
CC	-I- TISSUE SPECIFICITY: Expressed in brain and other tissues. Isoform		
CC	1 is neural-specific.		
CC	-I- SIMILARITY: Contains 23 ANK repeats.		
CC	-I- SIMILARITY: Contains 1 death domain.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; U13616; AAA64834.1; -		
DR	PUR; A55575; A55575.		
DR	HSP; P55273; IBI8.		
DR	Genew; HGNC:494; ANK3.		
DR	MIM; 600465; -		
DR	GO; GO:0006605; P;proteins targeting; NAS.		
DR	InterPro; IPR002110; ANK.		
DR	InterPro; IPR000488; Death.		
DR	InterPro; IPR000906; ZU5.		
DR	Pfam; PF00023; ank; 24.		
DR	Pfam; PF00531; death; 1.		
DR	Pfam; PF00791; ZU5; 1.		
DR	PRINTS; PR01415; ANKYRIN.		
DR	SMART; SM00248; ANK; 21.		
DR	SMART; SM00005; DEATH; 1.		
DR	SMART; SM00218; ZU5; 1.		
DR	PROSITE; PS50088; ANK_REPEAT; 21.		
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.		
DR	PROSITE; PS50017; DEATH_DOMAIN; 1.		
KW	Cytoskeleton; Alternative splicing; Repeat; ANK repeat.		
FT	REPEAT	73	102
FT	REPEAT	106	135
FT	REPEAT	139	168
FT	REPEAT	172	201
FT	REPEAT	203	230
FT	REPEAT	234	263
FT	REPEAT	267	296
FT	REPEAT	300	329
FT	REPEAT	333	362
FT	REPEAT	366	395

RESULT 9	
M3K7_HUMAN	
ID	M3K7_HUMAN
AC	043318; 043317; 043319; PRT; 606 AA.
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DE	Mitogen-activated protein kinase kinase 7 (BC 2.7.1.1-)
DE	(transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).
GN	MAP3K7 OR TAK1.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A. (ISOFORMS 1A; 1B AND 1C).
RC	TISSUE=Lung;
RX	MEDLINE=98153801; PubMed=9480845;
RA	Sakurai H., Shigemori N., Hasegawa K., Sugita T.;
RT	"TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an NF-kappa B-inducing kinase-independent mechanism.";
RL	Biochem. Biophys. Res. Commun. 243:545-549(1998).
RN	[2]
RP	SEQUENCE FROM N.A. (ISOFORM 1A).
RC	TISSUE=Uterus;
RX	MEDLINE=22388257; PubMed=12477932;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan K., Moore T., Max S.F., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Vallatton D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[3]
RP	INTERACTION WITH PELI1 AND PELI2.
RX	MEDLINE=22689054; PubMed=12804775;
RA	Jensen L.E., Whitehead A.S.;
RT	"Pellino2 activates the mitogen activated protein kinase pathway.";
RL	FEBS Lett. 545:199-202(2003).
RN	[4]
RP	INTERACTION WITH PELI3.
RX	MEDLINE=22756745; PubMed=12874243;
RA	Jensen L.E., Whitehead A.S.;
RT	"Pellino3, a novel member of the Pellino protein family, promotes activation of c-Jun and Elk-1 and may act as a scaffolding protein.";
RL	J. Immunol. 171:1500-1506(2003).
CC	-!- FUNCTION: Can phosphorylate and activate yet undefined MAPKs. Mediator of TGF-beta signal transduction. Stimulates NF-kappa-B activation.
CC	-!- SUBUNIT: Interacts with PELI1, PELI2 and PELI3.
CC	-!- ALTERNATIVE PRODUCTS:
CC	Event=Alternative splicing; Named isoforms=3;
CC	Name=1B;
CC	Isoid=043318-1; Sequence=Displayed;
CC	Name=1A;
CC	Isoid=043318-2; Sequence=VSP_004886;
CC	Name=1C;
CC	Isoid=043318-3; Sequence=VSP_004887; VSP_004888;
CC	-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. MAP

CC	kinase kinase kinase subfamily.
CC	-!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; ABC009357; BAA25026.1; -
DR	EMBL; ABC009356; BAA25025.1; -
DR	EMBL; ABC009358; BAA25027.2; -
DR	EMBL; BC017715; AAH17715.1; -
DR	PIR; JC5955; JC5955.
DR	PIR; JC5956; JC5956.
DR	HSP; P08631; IAD5.
DR	Genew; HGNC:16859; MAP3K7.
DR	MIM; 602614; -
DR	GO; GO:0004709; P:MAP kinase kinase activity; TAS.
DR	GO; GO:0007179; P:TGFbeta receptor signaling pathway; TAS.
DR	InterPro; IPR000719; Prot kinase.
DR	InterPro; IPR008271; Ser_Thr_kinase.
DR	InterPro; IPR001245; Tyr_kinase.
DR	Pfam; PF00069; pkinase; 1.
DR	PRINTS; PR00109; TYRKINASE.
DR	ProDom; PD000001; Prot kinase; 1.
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR	PROSITE; PS00108; PROTEIN KINASE_ST; 1.
DR	PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
KW	Transferase; Serine/threonine-protein kinase; ATP-binding; Alternative splicing.
KW	DOMAIN 8 14 POLY-SER.
FT	DOMAIN 36 291 PROTEIN KINASE.
FT	NP BIND 42 50 ATP (BY SIMILARITY).
FT	BINDING 63 63 ATP (BY SIMILARITY).
FT	ACT SITE 156 156 BY SIMILARITY.
FT	VARSPLIC 404 430 Missing (in isoform 1A).
FT	FTID=VSP_004886.
FT	FTID=VSP_004887.
FT	FTID=VSP_004888.
FT	FTID=VSP_004889.
FT	FTID=VSP_004890.
FT	FTID=VSP_004891.
FT	FTID=VSP_004892.
FT	FTID=VSP_004893.
FT	FTID=VSP_004894.
FT	FTID=VSP_004895.
FT	FTID=VSP_004896.
FT	FTID=VSP_004897.
FT	FTID=VSP_004898.
FT	FTID=VSP_004899.
FT	FTID=VSP_004900.
FT	FTID=VSP_004901.
FT	FTID=VSP_004902.
FT	FTID=VSP_004903.
FT	FTID=VSP_004904.
FT	FTID=VSP_004905.
FT	FTID=VSP_004906.
FT	FTID=VSP_004907.
FT	FTID=VSP_004908.
FT	FTID=VSP_004909.
FT	FTID=VSP_004910.
FT	FTID=VSP_004911.
FT	FTID=VSP_004912.
FT	FTID=VSP_004913.
FT	FTID=VSP_004914.
FT	FTID=VSP_004915.
FT	FTID=VSP_004916.
FT	FTID=VSP_004917.
FT	FTID=VSP_004918.
FT	FTID=VSP_004919.
FT	FTID=VSP_004920.
FT	FTID=VSP_004921.
FT	FTID=VSP_004922.
FT	FTID=VSP_004923.
FT	FTID=VSP_004924.
FT	FTID=VSP_004925.
FT	FTID=VSP_004926.
FT	FTID=VSP_004927.
FT	FTID=VSP_004928.
FT	FTID=VSP_004929.
FT	FTID=VSP_004930.
FT	FTID=VSP_004931.
FT	FTID=VSP_004932.
FT	FTID=VSP_004933.
FT	FTID=VSP_004934.
FT	FTID=VSP_004935.
FT	FTID=VSP_004936.
FT	FTID=VSP_004937.
FT	FTID=VSP_004938.
FT	FTID=VSP_004939.
FT	FTID=VSP_004940.
FT	FTID=VSP_004941.
FT	FTID=VSP_004942.
FT	FTID=VSP_004943.
FT	FTID=VSP_004944.
FT	FTID=VSP_004945.
FT	FTID=VSP_004946.
FT	FTID=VSP_004947.
FT	FTID=VSP_004948.
FT	FTID=VSP_004949.
FT	FTID=VSP_004950.
FT	FTID=VSP_004951.
FT	FTID=VSP_004952.
FT	FTID=VSP_004953.
FT	FTID=VSP_004954.
FT	FTID=VSP_004955.
FT	FTID=VSP_004956.
FT	FTID=VSP_004957.
FT	FTID=VSP_004958.
FT	FTID=VSP_004959.
FT	FTID=VSP_004960.
FT	FTID=VSP_004961.
FT	FTID=VSP_004962.
FT	FTID=VSP_004963.
FT	FTID=VSP_004964.
FT	FTID=VSP_004965.
FT	FTID=VSP_004966.
FT	FTID=VSP_004967.
FT	FTID=VSP_004968.
FT	FTID=VSP_004969.
FT	FTID=VSP_004970.
FT	FTID=VSP_004971.
FT	FTID=VSP_004972.
FT	FTID=VSP_004973.
FT	FTID=VSP_004974.
FT	FTID=VSP_004975.
FT	FTID=VSP_004976.
FT	FTID=VSP_004977.
FT	FTID=VSP_004978.
FT	FTID=VSP_004979.
FT	FTID=VSP_004980.
FT	FTID=VSP_004981.
FT	FTID=VSP_004982.
FT	FTID=VSP_004983.
FT	FTID=VSP_004984.
FT	FTID=VSP_004985.
FT	FTID=VSP_004986.
FT	FTID=VSP_004987.
FT	FTID=VSP_004988.
FT	FTID=VSP_004989.
FT	FTID=VSP_004990.
FT	FTID=VSP_004991.
FT	FTID=VSP_004992.
FT	FTID=VSP_004993.
FT	FTID=VSP_004994.
FT	FTID=VSP_004995.
FT	FTID=VSP_004996.
FT	FTID=VSP_004997.
FT	FTID=VSP_004998.
FT	FTID=VSP_004999.
FT	FTID=VSP_005000.
FT	FTID=VSP_005001.
FT	FTID=VSP_005002.
FT	FTID=VSP_005003.
FT	FTID=VSP_005004.
FT	FTID=VSP_005005.
FT	FTID=VSP_005006.
FT	FTID=VSP_005007.
FT	FTID=VSP_005008.
FT	FTID=VSP_005009.
FT	FTID=VSP_005010.
FT	FTID=VSP_005011.
FT	FTID=VSP_005012.
FT	FTID=VSP_005013.
FT	FTID=VSP_005014.
FT	FTID=VSP_005015.
FT	FTID=VSP_005016.
FT	FTID=VSP_005017.
FT	FTID=VSP_005018.
FT	FTID=VSP_005019.
FT	FTID=VSP_005020.
FT	FTID=VSP_005021.
FT	FTID=VSP_005022.
FT	FTID=VSP_005023.
FT	FTID=VSP_005024.
FT	FTID=VSP_005025.
FT	FTID=VSP_005026.
FT	FTID=VSP_005027.
FT	FTID=VSP_005028.
FT	FTID=VSP_005029.
FT	FTID=VSP_005030.
FT	FTID=VSP_005031.
FT	FTID=VSP_005032.
FT	FTID=VSP_005033.
FT	FTID=VSP_005034.
FT	FTID=VSP_005035.
FT	FTID=VSP_005036.
FT	FTID=VSP_005037.
FT	FTID=VSP_005038.
FT	FTID=VSP_005039.
FT	FTID=VSP_005040.
FT	FTID=VSP_005041.
FT	FTID=VSP_005042.
FT	FTID=VSP_005043.
FT	FTID=VSP_005044.
FT	FTID=VSP_005045.
FT	FTID=VSP_005046.
FT	FTID=VSP_005047.
FT	FTID=VSP_005048.
FT	FTID=VSP_005049.
FT	FTID=VSP_005050.
FT	FTID=VSP_005051.
FT	FTID=VSP_005052.
FT	FTID=VSP_005053.
FT	FTID=VSP_005054.
FT	FTID=VSP_005055.
FT	FTID=VSP_005056.
FT	FTID=VSP_005057.
FT	FTID=VSP_005058.
FT	FTID=VSP_005059.
FT	FTID=VSP_005060.
FT	FTID=VSP_005061.
FT	FTID=VSP_005062.
FT	FTID=VSP_005063.
FT	FTID=VSP_005064.
FT	FTID=VSP_005065.
FT	FTID=VSP_005066.
FT	FTID=VSP_005067.
FT	FTID=VSP_005068.
FT	FTID=VSP_005069.
FT	FTID=VSP_005070.
FT	FTID=VSP_005071.
FT	FTID=VSP_005072.
FT	FTID=VSP_005073.
FT	FTID=VSP_005074.
FT	FTID=VSP_005075.
FT	FTID=VSP_005076.
FT	FTID=VSP_005077.
FT	FTID=VSP_005078.
FT	FTID=VSP_005079.
FT	FTID=VSP_005080.
FT	FTID=VSP_005081.
FT	FTID=VSP_005082.
FT	FTID=VSP_005083.
FT	FTID=VSP_005084.
FT	FTID=VSP_005085.
FT	FTID=VSP_005086.
FT	FTID=VSP_005087.
FT	FTID=VSP_005088.
FT	FTID=VSP_005089.
FT	FTID=VSP_005090.
FT	FTID=VSP_005091.
FT	FTID=VSP_005092.
FT	FTID=VSP_005093.
FT	FTID=VSP_005094.
FT	FTID=VSP_005095.
FT	FTID=VSP_005096.
FT	FTID=VSP_005097.
FT	FTID=VSP_005098.
FT	FTID=VSP_005099.
FT	FTID=VSP_005100.
FT	FTID=VSP_005101.
FT	FTID=VSP_005102.
FT	FTID=VSP_005103.
FT	FTID=VSP_005104.
FT	FTID=VSP_005105.
FT	FTID=VSP_005106.
FT	FTID=VSP_005107.
FT	FTID=VSP_005108.
FT	FTID=VSP_005109.
FT	FTID=VSP_005110.
FT	FTID=VSP_005111.
FT	FTID=VSP_005112.
FT	FTID=VSP_005113.
FT	FTID=VSP_005114.
FT	FTID=VSP_005115.
FT	FTID=VSP_005116.
FT	FTID=VSP_005117.
FT	FTID=VSP_005118.
FT	FTID=VSP_005119.
FT	FTID=VSP_005120.
FT	FTID=VSP_005121.
FT	FTID=VSP_005122.
FT	FTID=VSP_005123.
FT	FTID=VSP_005124.
FT	FTID=VSP_005125.
FT	FTID=VSP_005126.
FT	FTID=VSP_005127.
FT	FTID=VSP_005128.
FT	FTID=VSP_005129.
FT	FTID=VSP_005130.
FT	FTID=VSP_005131.
FT	FTID=VSP_005132.
FT	FTID=VSP_005133.
FT	FTID=VSP_005134.
FT	FTID=VSP_005135.
FT	FTID=VSP_005136.
FT	FTID=VSP_005137.
FT	FTID=VSP_005138.
FT	FTID=VSP_005139.
FT	FTID=VSP_005140.
FT	FTID=VSP_005141.
FT	FTID=VSP_005142.
FT	FTID=VSP_005143.
FT	FTID=VSP_005144.
FT	FTID=VSP_005145.
FT	FTID=VSP_005146.
FT	FTID=VSP_005147.
FT	FTID=VSP_005148.
FT	FTID=VSP_005149.
FT	FTID=VSP_005150.
FT	FTID=VSP_005151.
FT	FTID=VSP_005152.
FT	FTID=VSP_005153.
FT	FTID=VSP_005154.
FT	FTID=VSP_005155.
FT	FTID=VSP_005156.
FT	FTID=VSP_005157.
FT	FTID=VSP_005158.
FT	FTID=VSP_005159.
FT	FTID=VSP_005160.
FT	FTID=VSP_005161.
FT	FTID=VSP_005162.
FT	FTID=VSP_005163.
FT	FTID=VSP_005164.
FT	FTID=VSP_005165.
FT	FTID=VSP_005166.
FT	FTID=VSP_005167.
FT	FTID=VSP_005168.
FT	FTID=VSP_005169.
FT	FTID=VSP_005170.
FT	FTID=VSP_005171.
FT	FTID=VSP_005172.
FT	FTID=VSP_005173.
FT	FTID=VSP_005174.
FT	FTID=VSP_005175.
FT	FTID=VSP_005176.
FT	FTID=VSP_005177.
FT	FTID=VSP_005178.
FT	FTID=VSP_005179.
FT	FTID=VSP_005180.
FT	FTID=VSP_005181.
FT	FTID=VSP_005182.
FT	FTID=VSP_005183.
FT	FTID=VSP_005184.
FT	FTID=VSP_005185.
FT	FTID=VSP_005186.
FT	FTID=VSP_005187.
FT	FTID=VSP_005188.
FT	FTID=VSP_005189.
FT	FTID=VSP_005190.
FT	FTID=VSP_005191.
FT	FTID=VSP_005192.
FT	FTID=VSP_005193.
FT	FTID=VSP_005194.
FT	FTID=VSP_005195.
FT	FTID=VSP_005196.
FT	FTID=VSP_005197.
FT	FTID=VSP_005198.
FT	FTID=VSP_005199.
FT	FTID=VSP_005200.
FT	FTID=VSP_005201.
FT	FTID=VSP_005202.
FT	FTID=VSP_005203.
FT	FTID=VSP_005204.
FT	FTID=VSP_005205.
FT	FTID=VSP_005206.
FT	FTID=VSP_005207.
FT	FTID=VSP_005208.
FT	FTID=VSP_005209.
FT	FTID=VSP_005210.
FT	FTID=VSP_005211.
FT	FTID=VSP_005212.
FT	FTID=VSP_005213.
FT	FTID=VSP_005214.
FT	FTID=VSP_005215.
FT	FTID=VSP_005216.
FT	FTID=VSP_005217.
FT	FTID=VSP_005218.
FT	FTID

SEQUENCE FROM N.A.
 RC STRAIN=cv, Columbia; TISSUE=Seedling;
 RX MEDLINE=93161417; PubMed=9431946;
 RA Kieber J.J., Rothenberg M., Roman G., Feldmann K.A., Ecker J.R.;
 RT "CTR1, a negative regulator of the ethylene response pathway in
 Arabidopsis, encodes a member of the raf family of protein kinases."
 RL Cell 72:427-441(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv, Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 Nakazaki N., Naruo K., Okumura S., Shinjo S., Takeuchi C., Wada T.,
 Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 Habermann K., Murray J., Johnson D., Rohlfing T., Nelson R.,
 Stabenking T., Peplin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 Balter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strowmatt C.,
 Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
 Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 Kirchhoff K., Roth K., King L., Bahret A., Miller B., Marra M.A.,
 Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
 Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 Langham S.A., McQuillagh B., Robben J., Grynoprez B., Zimmermann W.,
 Ramsberger U., Wedler H., Balke K., Wedler E., Peters S.,
 van Staveren M., Dirkee W., Mooijman P., Klein Lankhorst R.,
 Weissenegger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
 Feldpausch M., Lamberth S., Villarroel R., Giesen J., Ardiles W.,
 Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana."
 RL Nature 408:823-826(2000).
 CC -!- FUNCTION: Acts as a negative regulator in the ethylene response
 pathway.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- TISSUE SPECIFICITY: Expressed in both seedlings and adult plants.
 CC -!- MISCELLANEOUS: CTR1 mutants display ethylene-treated phenotypes,
 resulting in plants with small, unexpanded leaves and whose seed
 cotelodion growth is impaired.
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. RAP
 subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L08789; AAA32779.1; -;
 CC EMBL; L08790; AAA32780.1; -;
 CC EMBL; AL162506; CAB82938.1; -;
 CC PIR; T48400; T48400.
 CC InterPro; IPR000719; Prot kinase.
 CC InterPro; IPR008271; Ser Thr pkin AS.
 CC InterPro; IPR002290; Ser Thr pkinase.
 CC InterPro; IPR001245; Tyr_pkinase.
 CC Pfam; PF00069; pkinase; 1.
 CC PRINTS; PR00109; TYRKINASE.
 CC ProDom; PD000001; Prot kinase; 1.
 CC SMART; SM00220; S_TKc; 1.
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
 CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
 CC DOMAIN 65 69 POLY-GLY.

FT DOMAIN 135 141 POLY-GLY.
 FT DOMAIN 551 809 PROTEIN KINASE.
 FT NP_BIND 557 565 ATP (BY SIMILARITY).
 FT BINDING 578 578 ATP (BY SIMILARITY).
 FT ACT_SITE 676 596 BY SIMILARITY.
 FT MUTAGEN 596 596 E->K: IN CTR1-4; EXHIBITS ETHYLENE-
 TREATED PHENOTYPE.
 FT MUTAGEN 694 694 D->E: IN CTR1-1; EXHIBITS ETHYLENE-
 TREATED PHENOTYPE.
 SQ SEQUENCE 821 AA; 90306 MW; 29223DCDC0CC15BC CRC64;
 Query Match 9.2%; Score 404.5; DB 1; Length 821;
 Best Local Similarity 26.1%; Pred. No. 2.9e-17;
 Matches 139; Conservative 90; Mismatches 206; Indels 97; Gaps 20;
 QY 273 P L H A C Y N G K F V A K E I I Q I S G T E ----- S L T K E N I F S T A F H S A C T Y G K S I D - 320
 DB 329 P M W K E I D G L K I F K V V V P I G S L V G L C H R A L L F K V L A D I I D L P C R T A K G C K Y C N R D D A 388
 QY 321 --- L V K F L L D Q N V I N H O G R G H ----- T G L H S A C Y G H I R L V Q - F L L D N G A D M N L 368
 DB 389 A S C L V R F G L D R E Y L - V D L V G K P G H L W E P D S L L N G P S S I S I S S P L R F P R K F V E P A V D F R L 447
 QY 369 V A C D P S R S S G E D E Q T C L M W A Y E K G H D A L V T L L K H Y K R P Q D E L P C N E Y S Q P G G S Y V - - 426
 DB 448 L A ----- K O Y F S D S Q S - L N L V F P A S D D M G F S M F H ----- R Q Y D N F G G E N D A L A E 491
 QY 427 --- S V P S P L G I K S M T K E K A D I L L R A G L ----- P S H F L Q L S ----- E 462
 DB 492 N G G G S L P - P S A N M P P Q N M W R A S N I E A A P M N A P P I S Q P V P N R A N R E L G D G D M D I P W C D 550
 QY 463 I E P H E I I G S G S G K Y K G E C R N K I V A I K Y R A N T Y C S K S D Y D M E C R E Y S I L C O L N H P C V I 522
 DB 551 L N T K E I G A G S G T V H R A E W H G S D V A V K L I M E Q D P A E R - V N E F L R E V A I M K R L H P N I V 609
 QY 523 Q F V G A C I N D P S Q F A I V T Q I S G S L F S L L H E Q - - K R I L D L Q S K L I I A V D V A K G M E Y L H N L 580
 DB 610 L F W G A - V T O P P N L S I V T E Y L S R G S L Y R L L K S G A R E Q L D E R R L S M A Y D V A K G M Y L H N R 668
 QY 581 T Q P I I H R D L N S H I L Y E D G H A V A D F G E S R ----- F L Q S L D E D N M T Q P G N L R W A P E V 635
 DB 669 N P P I V H R D L K S P N L L V D K Y T V K V C D F G L S R L K A S T F L S S ----- K S A A G T P E W M A P E V 722
 QY 636 F T Q C T R Y T I K A D V E S V A L C L W E I L T G E I P F A H L K P A A A A A D M A Y H H I P P I G Y S I P K P I S 695
 DB 723 L R D - E F S N E K S D V Y S F G V I L W E L A T I Q P W G N L N A Q V A A V G F K C K L E I P R N L N P Q V A 781
 QY 696 S L L I R G W N A C P E R P E F S E V M K L E C L N I E L M S P A S S N S G S L S P S S S S D 747
 DB 782 A I I E G C W T N E P W K R P S F A T I ----- M D L L R P L I K S ----- A V P P P N R S D 820
 RESULT 12
 HT16 HYDAT
 ID HT16 HYDAT STANDARD; PRT; 757 AA.
 AC P53356;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase HTK16 (EC 2.7.1.112).
 GN HTK16.
 OS Hydra attenuata (Hydra) (Hydra vulgaris).
 CC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Anthomedusae;
 CC Hydridae; Hydra.
 CC NCBI_TaxID=6087;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Irvine;
 RX MEDLINE=94181280; PubMed=8134129;
 RA Chan T.A., Chu C.A., Rauen K.A., Krother M., Tatarewicz S.M.,
 RA Steele R.E.;
 RT "Identification of a gene encoding a novel protein-tyrosine kinase
 containing SH2 domains and ankyrin-like repeats."

RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyx S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.P., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=98409636; PubMed=9736715;
RA Delcommenne M., Tan C., Gray V., Rue L., Woodgett J., Dedhar S.;
RT "Phosphoinositide-3-OH kinase-dependent regulation of glycogen
RT synthase kinase 3 and protein kinase B/AKT by the integrin-linked
RT kinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:11211-11216(1998).
RN [6]
RP REVIEW.
RX MEDLINE=20179495; PubMed=10712922;
RA Dedhar S.;
RT "Cell-substrate interactions and signaling through ILK.";
RL Curr. Opin. Cell Biol. 12:250-256(2000).
CC -|- FUNCTION: RECEPTOR-PROXIMAL PROTEIN KINASE REGULATING INTEGRIN-
CC MEDIATED SIGNAL TRANSDUCTION. MAY ACT AS A MEDIATOR OF INSIDE-OUT
CC INTEGRIN SIGNALING. FOCAL ADHESION PROTEIN PART OF THE COMPLEX
CC ILK-FINCH. THIS COMPLEX IS CONSIDERED TO BE ONE OF THE CONVERGENCE
CC POINTS OF INTEGRIN- AND GROWTH FACTOR-SIGNALING PATHWAY. COULD BE
CC IMPLICATED IN MEDIATING CELL ARCHITECTURE, ADHESION TO INTEGRIN
CC SUBSTRATES AND ANCHORAGE-DEPENDENT GROWTH IN EPITHELIAL CELLS.
CC PHOSPHORYLATES BETA-1 AND BETA-3 INTEGRIN SUBUNIT ON SERINE AND
CC THREONINE RESIDUES, BUT ALSO AKT1 AND GSK3B.
CC -|- ENZYME REGULATION: STIMULATED RAPIDLY BUT TRANSIENTLY BY BOTH CELL
CC FIBRONECTIN INTERACTIONS, AS WELL AS BY INSULIN, IN A PI(3)K-
CC DEPENDENT MANNER, LIKELY VIA THE BINDING OF PTDINS(3,4,5)P3 WITH A
CC PH-LIKE DOMAIN OF ILK1.
CC -|- SUBUNIT: INTERACTS WITH CYTOPLASMIC DOMAIN OF BETA 1 SUBUNIT OF
CC INTEGRIN. COULD ALSO INTERACTS WITH BETA 2, BETA 3 AND/OR BETA 5
CC SUBUNIT OF INTEGRIN. ALSO INTERACTS WITH PINCH AND PARVINS.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART FOLLOWED BY SKELETAL
CC MUSCLE, PANCREAS AND KIDNEY. WEAKLY EXPRESSED IN PLACENTA, LUNG
CC AND LIVER.
CC -|- DOMAIN: A PH-LIKE DOMAIN IS INVOLVED IN PHOSPHATIDYLINOSITOL
CC PHOSPHATE BINDING.
CC -|- PTM: Autophosphorylated on serine residues.
CC -|- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -|- SIMILARITY: Contains 3 ANK repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U40282; AAC16892.1; -;
CC EMBL: AF244139; BAB74449.1; -;
CC EMBL: A404847; CAB99253.1; -;
CC EMBL: BC001554; AAB01554.1; -;

DR HSP; Q00420; IAWC.
DR Genew; HGNC:6040; ILK.
DR MIM; 602366; -;
DR GO; GO:0004674; F-protein serine/threonine kinase activity; TAS.
DR GO; GO:0007160; P-cell-matrix adhesion; TAS.
DR GO; GO:0007229; P-integrin-mediated signaling pathway; TAS.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr_kin_AS.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00023; ank; 3.
DR Pfam; PF00069; kinase; 1.
DR PRINTS; PRO1415; TYRKINASE.
DR PRINTS; PRO1019; TYRKINASE.
DR SMART; SM00248; ANK; 3.
DR SMART; PD000001; Prot kinase; 1.
DR PROSITE; PS50297; ANK_REPEAT; 1.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Repeat; ANK repeat.
FT REPEAT 33 82 ANK 1.
FT REPEAT 66 95 ANK 2.
FT REPEAT 99 128 ANK 3.
FT DOMAIN 180 212 PH-LIKE.
FT DOMAIN 193 446 PROTEIN_KINASE.
FT NP_BIND 199 207 ATP (BY SIMILARITY).
FT BINDING 220 220 ATP (BY SIMILARITY).
FT MUTAGEN 359 359 E->K: INACTIVATION OF ILK.
SQ SEQUENCE 452 AA; 51419 MW; E37DC2AD5311A1C2 CRC64;
Query Match 9.1%; Score 400.5; DB 1; Length 452;
Best Local Similarity 23.0%; Pred. No. 2,3e-17;
Matches 128; Conservative 94; Mismatches 174; Indels 161; Gaps 15;
Qy 192 DVNVSEVGDRLPHLASAKGFLNIAKLMEBSKADVAQDNEDHVPHFCSFGHHDIV 251
Db 26 DLNQGDHGFSLHWACREGSAVVVEMLIMRGAR--INVMRGDTPHLAASHGHHDIV 83
Qy 252 KYLLQSDLEVQHVNVIGDTPHLACVNGKEVAKETIQTSGTSLTKENIFSETAFHS 311
Db 84 QKLLQYKADI--NAVNEHGVNVEFLHYACFWGQDQVAEDLVA----- 121
Qy 312 ACTYGSIDLVKFLLDQNVININHQGRDHTGLHSACYHGHRLVQFLDNGADNLVAC 371
Db 122 -----NGALVSI--- 128
Qy 372 DPRSSEGEDEQCLMWAVEKGHDAIVTLKHYKEPQDELPCNEYSPGGGSGYVSPSP 431
Db 129 -----CNRYGEMFVD-----KAKAP 143
Qy 432 LGKIKSMTEKADILLRL-----AGLPSHFHLQLSEIEFHEIIGSGS 473
Db 144 LRELLRERAEKMGONLNRIPYKDTFWKGTTRTPRNGTLNKGSGIDFKQLNFLTCLNENH 203
Qy 474 FGKVKRCRNKIVAKRYRANTYCKSDVDMFCRENSILCOLNHPCVIQFVGLCLNDPS 533
Db 204 SGLWKGWRQGNDIVVKLVKVRDWSRTRSRD--FNECPRLRIFSHFNVLFPVIGACQSP 262
Qy 534 QF-AIVTQVIGSGISFLSLHEQKRLDLSKLI--TAVDVAKGMEYLNHLTQPIIHRD--LN 590
Db 263 PPHTLTHMPYGSGLYNVLHGTNFEVDSQAVKFDLDMARGMAFLHTL--EPLPRHALN 321
Qy 591 SHNILEYDGHAVADFESRFLQSLDSDNMTKQ--PGNL-----RWNAPEVFTQCTRYTK- 645
Db 322 SRSMVIDEDMTARI-----SNADVKTFSQCPGRYAPAWVAPEALCKPEDTNR 371
Qy 646 -ADVESVALCLWEILTGETPPFAHLKAPAAADMAVHHTRPPIGYSIKPIISLILRGNA 704
Db 372 SADMSFAVILLWELVTREVPFADLNMEIGMKVALEBURPTIPPGISPHVCKLMKICWNE 431

QY 705 CPEGRPEFSEVVMKLE 721
Db 432 DPAKRPFDMDIVPILEK 448

RESULT 14
ILK_MOUSE
ID ILK_MOUSE STANDARD; - PRT; 452 AA.
AC 055222;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin-linked protein kinase (EC 2.7.1.1-).
GN ILK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; Tissue=Heart;
RX MEDLINE=98031580; PubMed=9366252;
RA Li F., Liu J., Mayne R., Wu C.;
RT "Identification and characterization of a mouse protein kinase that is
RL highly homologous to human integrin-linked kinase.";
RL Biochim. Biophys. Acta 1359:215-220(1997).
CC !- FUNCTION: RECEPTOR-PROXIMAL PROTEIN KINASE REGULATING INTEGRIN-
CC MEDIATED SIGNAL TRANSDUCTION. MAY ACT AS A MEDIATOR OF INSIDE-OUT
CC INTEGRIN SIGNALING. FOCAL ADHESION PROTEIN PART OF THE COMPLEX
CC ILK-PINCH. THIS COMPLEX IS CONSIDERED TO BE ONE OF THE CONVERGENCE
CC POINTS OF INTEGRIN- AND GROWTH FACTOR-SIGNALING PATHWAY. COULD BE
CC IMPLICATED IN MEDIATING CELL ARCHITECTURE, ADHESION TO INTEGRIN
CC SUBSTRATES AND ANCHORAGE-DEPENDENT GROWTH IN EPITHELIAL CELLS.
CC PHOSPHORYLATES BETA-1 AND BETA-3 INTEGRIN SUBUNIT ON SERINE AND
CC THREONINE RESIDUES, BUT ALSO AKT1 AND GSK3B (BY SIMILARITY).
CC !- SUBUNIT: INTERACTS WITH CYTOPLASMIC DOMAIN OF BETA 1 SUBUNIT OF
CC INTEGRIN. COULD ALSO INTERACTS WITH BETA 2, BETA 3 AND/OR BETA 5
CC SUBUNIT OF INTEGRIN. ALSO INTERACTS WITH PINCH AND PARVINS (BY
CC similarity).
CC !- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC !- TISSUE SPECIFICITY: Highly expressed in lung, heart, kidney,
CC liver, brain, spleen and skeletal muscle. Weakly expressed in
CC testis.
CC !- DOMAIN: A PH-LIKE DOMAIN IS INVOLVED IN PHOSPHATIDYLINOSITOL
CC PHOSPHATE BINDING (BY SIMILARITY).
CC !- PTM: Autophosphorylated on serine residues (By similarity).
CC !- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC !- SIMILARITY: Contains 3 ANK repeats.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U94479; AAB94646.1; -
CC HSSP; Q00420; IAWC.
CC MGD; MG1:1195267; ILK.
CC GO; GO:0005925; C:focal adhesion; IDA.
CC GO; GO:0005515; F:protein binding; IPI.
CC GO; GO:0045197; P:establishment and/or maintenance of epithel. . .; IMP.
CC DR InterPro; IPR002110; ANK.
CC DR InterPro; IPR000719; Prot. kinase.
CC DR InterPro; IPR008271; Ser_thr_pkin_AS.
CC DR InterPro; IPR001245; Tyr_pkinase.
CC DR Pfam; PF00023; ank; 3.
CC DR Pfam; PF00069; pkinase; 1.
CC DR PRINTS; PR01415; ANKYRIN.
CC DR PRINTS; PR00109; TYRKINASE.
CC DR ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00248; ANK; 3.

DR PROSITE; PS50297; ANK REP REGION; 1.
DR PROSITE; PS50088; ANK REPEAT; 3.
DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN KINASE_ST; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW phosphorylation; Repeat; ANK repeat.
FT REPEAT 33 62 ANK 1.
FT REPEAT 66 95 ANK 2.
FT REPEAT 99 128 ANK 3.
FT DOMAIN 180 212 PH-LIKE.
FT DOMAIN 193 446 PROTEIN KINASE.
FT NP_BIND 199 207 ATP (BY SIMILARITY).
FT BINDING 220 220 ATP (BY SIMILARITY).
SQ SEQUENCE 452 AA; 51347 MW; E0BB30C52F6865E CRC64;

Query Match 9.0%; Score 396; DB 1; Length 452;
Best Local Similarity 23.6%; Pred. No. 4.4e-17;
Matches 126; Conservative 97; Mismatches 185; Indels 126; Gaps 15;

QY 215 IAKLMEEGSKADVNAQDNEDHVPFHFCRSFGRGHHDIVKYLQSDLEQVPHVNIYGDTP 274
Db 14 VAVRLWLDNTENDLNQDDHGFSPHWCACREGSAVEMLMRGARI--NVMNRGDDTP 71
QY 275 HLACYNGKFEVAKELIIOISGTSLETKENIFSETAFHSACTYKSIDLVKFLDDQVIN 334
Db 72 HLAASHGHRRDIVOKLIQ-----YKADTNVAVN-- 97
QY 335 HQGRDGHGTGLHSAHYGHIRLVQFLDNGADNMNLVACDPSSRSSEKEDQOTCLMAYEKGH 394
Db 98 ---EHGNVPLHYACFWGQDQVAEDLVANGALVSI----- 128
QY 395 DAIVTLKHKYRPQDELPCNEYSQPGGDSYVSVPGLKIKSMTEKXADILLR----- 449
Db 129 -----CNKYGEMPEVD-----KAKAPLRELLREAREKNGQNLNRIPKYD 166
QY 450 -----AGLPSPHFHLQLSIEFHEIIGSGFGKVGKVCNKRKIVAKRYRANT 496
Db 167 TFWKGTTRTPRNGTLNKGSGIDFKQLNFKLNENHSGELWKGWQGNDIWVKLVKRD 226
QY 497 YCSKSDVDMFCREVSLICQINHPCVIQFVAGACLNPSQF-AIVTOYISGSLPSLLHQK 555
Db 227 WSTRKSRD-FNEECPRLRIFSHPNVLPVLGACQAPPAHPHTLTITHTMPYGSLYNVLHEGT 285
QY 556 RILDLSKLI-IAVDYAKGMEVHLNLTQPIIHRD-LNSHNILLYEDGHAVVADFGESRFL 613
Db 286 NFVVDQSQAVKFAEDWARGMAFLHTD-EPLIPHALNSRSMIDEDMTARI----- 335
QY 614 QSLDEDNMTKQ-EGNL---RWMAPEVFTQTRYTK--ADVFYSALCWELITGTIPFAH 667
Db 336 -SMADVKSFCQGRVYAPAWVAPEALQKKPEDTNRSSADWMSFAVLLMELVTRVFPAD 394
QY 668 LKPAARAAADWYHHIPPGYISIPKPISSILLINGWNAECPEGRPEFSEVVMKLE 721
Db 395 LSNMEIGMKVALGLRPTVPPGISPHVCKMKICMKNEDPAKRPKFDIMVPILEK 448

RESULT 15
ILK2_HUMAN
ID ILK2_HUMAN STANDARD; PRT; 452 AA.
AC P57043;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin-linked protein kinase 2 (EC 2.7.1.1-). (ILK-2).
GN ILK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20332243; PubMed=10871859;

RA Janji B., Melchior C., Vallar L., Kieffer N.;
 RT "Cloning of an isoform of integrin-linked kinase (ILK) that is
 RT upregulated in HT-144 melanoma cells following TGF-beta1
 RT stimulation.";
 RL Oncogene 19:3069-3077(2000).
 CC -!- FUNCTION: RECEPTOR-PROXIMAL PROTEIN KINASE REGULATING INTEGRIN-
 CC MEDIATED SIGNAL TRANSDUCTION. MAY ACT AS A MEDIATOR OF INSIDE-OUT
 CC INTEGRIN SIGNALING. FOCAL ADHESION PROTEIN PART OF THE COMPLEX
 CC ILK-PINCH. THIS COMPLEX IS CONSIDERED TO BE ONE OF THE CONVERGENCE
 CC POINTS OF INTEGRIN- AND GROWTH FACTOR-SIGNALING PATHWAY. COULD BE
 CC IMPLICATED IN MEDIATING CELL ARCHITECTURE, ADHESION TO INTEGRIN
 CC SUBSTRATES AND ANCHORAGE-DEPENDENT GROWTH IN EPITHELIAL CELLS.
 CC PHOSPHORYLATES BETA-1 AND BETA-3 INTEGRIN SUBUNIT ON SERINE AND
 CC THREONINE RESIDUES, BUT ALSO AKT1 AND GSK3B (BY SIMILARITY).
 CC -!- SUBUNIT: INTERACTS WITH CYTOPLASMIC DOMAIN OF BETA 1 SUBUNIT OF
 CC INTEGRIN. COULD ALSO INTERACTS WITH BETA 2, BETA 3 AND/OR BETA 5
 CC SUBUNIT OF INTEGRIN. ALSO INTERACTS WITH PINCH AND PARVINS (BY
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in some highly invasive tumor cell
 CC lines but not in normal tissues.
 CC -!- INDUCTION: By TGF-beta1.
 CC -!- DOMAIN: A PH-LIKE DOMAIN IS INVOLVED IN PHOSPHATIDYLINOSITOL
 CC PHOSPHATE BINDING (BY SIMILARITY).
 CC -!- PTM: Autophosphorylated on serine residues.
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC -!- SIMILARITY: Contains 3 ANK repeats.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ277481; CAB94832.1; --
 CC HSSP; Q00420; 1AWC.
 CC GO; GO:0005737; C:cytoplasm; NAS.
 CC GO; GO:0005524; F:ATP binding; NAS.
 CC GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
 CC GO; GO:0008283; P:cell proliferation; NAS.
 CC GO; GO:0007160; P:cell-matrix adhesion; NAS.
 CC GO; GO:0007229; P:integrin-mediated signaling pathway; NAS.
 CC GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
 CC InterPro; IPR002110; ANK.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR008271; Ser_thr_pkin_AS.
 CC Pfam; PF00023; ank; 3.
 CC Pfam; PF00069; pkinase; 1.
 CC PRINTS; PR01415; ANKYRIN.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00248; ANK_3_kinase; 1.
 CC PROSITE; PS00297; ANK_REPEAT_REGION; 1.
 CC PROSITE; PS00088; ANK_REPEAT; 3.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 CC PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 CC Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Repeat; ANK repeat.
 FT REPEAT 33 62 ANK 1.
 FT REPEAT 66 95 ANK 2.
 FT REPEAT 99 128 ANK 3.
 FT DOMAIN 180 212 PH-LIKE.
 FT DOMAIN 193 446 PROTEIN_KINASE.
 FT NP_BIND 199 207 ATP (BY SIMILARITY).
 FT BINDING 220 220 ATP (BY SIMILARITY).
 SQ SEQUENCE 452 AA; 51393 MW; BF520214E373DC88 CRC64;

Query Match 8.9%; Score 390.5; DB 1; Length 452;
 Best Local Similarity 22.8%; Pred. No. 9,5e-17;
 Matches 127; Conservative 94; Mismatches 175; Indels 161; Gaps 15;

QY 192 DVNVSGEVGDEPLHLASAKGLINIAKLMEEGSKADVNAQDNEDHVLHFCSEFGHHDIV 251
 Db 26 DLNQGDDHGFSLHWACREGRSAVVEMLMIEGAR--INVMRGDDTPLHLAASHGHHRDIV 83
 QY 252 KYLLQSDLEVQPHVNIYDTPHLHLACVNGKFEVAKETIIQISGTESLTKEINIFSETAFHS 311
 Db 84 OKLLOKADI--NAVNEHGNVPLHYACFWGQDVAEDLVA----- 121
 QY 312 ACTYKSIDLVKFLDDQNVININHQGRDHTGLHSACVHGHIRLVQVFLDNGADMLVAC 371
 Db 122 -----NGALVSI--- 128
 QY 372 DPSRSGEKDEQTCMLWAYEKGHDAIVTLKHYYRQDELPCEYEQSGDGSVSVSP 431
 Db 129 -----CNKYGEMPV-----KAKAP 143
 QY 432 LGKIKSMTEKADILLR-----AGLPSEHFLQLSETEFHEIIGSGS 473
 Db 144 LRELLREAEKMGQNLNRIPYKDTFWKGTTRPRNGTLNKGSGIDFKQLNFLAKLNENH 203
 QY 474 FGKYVKGRCRNKIVAIAKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLNDRPS 533
 Db 204 SGELWKGKRWQSNDIVVKVLKVRDMSTRKSRD-FNEECPRLRIFSHPNVLPVLGACQAPPA 262
 QY 534 QF-AIVTOYISGSLFSLHSEOKRILD-OSKLI-IADVAVKMEYLNLTQPIIHRD-LN 590
 Db 263 PHPTLIITHWPGYSLYVNLHSGTNFVDQSQAVKALDMARGMAFLHTL-EPLIPRALN 321
 QY 591 SHNILLYEDGHAVVADFGESRFLOSLEDNNMTKQ-PGNL---RWMAPSEVFTQCTRYTIK- 645
 Db 322 SRSVMDEDMTARI-----SMADVKEFQCPGRMVAPAMVAPEALQKKPEDTNRR 371
 QY 646 -ADVSEYALCLWEILLTGPIPAHLKPAADAAADMAHYHRRPPIGYSIKPISILLIRGWA 704
 Db 372 SADMMSPAVLLWELVTREVPFADLSNMEIGMKVALEGLRPTIPPGISPHVCKLMKICMNE 431
 QY 705 CPEGRPEFSEVVMKLEE 721
 Db 432 DPAKRSKFTDMIVPILK 448

Search completed: September 2, 2004, 14:21:46
 Job time : 39 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2004, 19:33:36 ; Search time 6441.43 Seconds
(without alignments)
16855.624 Million cell updates/sec

Title: US-10-626-173-3

Perfect score: 2505

Sequence: 1 aggggaattataacttag.....gcagctttgaggacagcgc 2505

Scoring table: IDENTITY_NUC

Gapop 10.0 , Capext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rat.*
- 36: em.htg.rod.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2505	100.0	2505	6	BD269560	Novel CAR
2	2505	100.0	2505	6	AR269112	Sequence
3	2505	100.0	2505	6	AR437546	Sequence
4	2505	100.0	2505	6	AX056437	Sequence
5	2505	100.0	2890	9	AY303691	Homo sapi
6	2505	100.0	3001	9	AF116826	Homo sapi
7	2505	100.0	3025	6	BD269559	Novel CAR
8	2505	100.0	3025	6	AR269111	Sequence
9	2505	100.0	3025	6	AR437545	Sequence
10	1892.2	75.5	2505	6	BD269562	Novel CAR
11	1892.2	75.5	2505	6	AR269114	Sequence
12	1892.2	75.5	2505	6	AR437548	Sequence
13	1892.2	75.5	2982	10	AY303692	Novel CAR
14	1892.2	75.5	3026	6	BD269561	Novel CAR
15	1892.2	75.5	3026	6	AR269113	Sequence
16	1892.2	75.5	3026	6	AR437547	Sequence
17	1730.6	69.1	7507	9	HSN807040	Homo sapi
18	166.4	18.6	568	6	AX525654	Sequence
19	199.4	8.0	150272	2	AC073887	Homo sapi
20	199.4	8.0	179582	9	AC093158	Homo sapi
21	196.2	7.8	66066	2	AC026566	Homo sapi
22	177.6	7.1	112358	9	AC105271	Homo sapi
23	166.6	6.7	197553	10	AC124174	Mus muscu
24	165	6.6	269035	2	AC126082	Rattus no
25	185	6.6	276946	2	AC115396	Rattus no
26	153.4	6.1	150272	2	AC073887	Homo sapi
27	153	6.1	146018	2	AC018404	Homo sapi
28	149.4	6.0	66066	2	AC026566	Homo sapi
29	138.2	5.5	258378	2	AC105566	Rattus no
30	110.4	4.4	78331	9	AC098692	Homo sapi
31	107.2	4.3	60037	2	AC100314	Mus muscu
32	107.2	4.3	254819	2	AC144762	Mus muscu
33	106.4	4.2	1672	8	AY096470	Arabidops
34	106.4	4.2	2122	8	AY070086	Arabidops
35	101.4	4.0	28196	9	AC119672	Homo sapi
36	100.2	4.0	1704	6	AX652967	Sequence
37	100.2	4.0	2263	8	AK072690	Oryza sat
38	99.4	4.0	1748	8	AY128938	Arabidops
39	99.4	4.0	2104	8	AY093017	Arabidops
40	98	3.9	1662	6	AX506319	Sequence
41	94.8	3.8	1961	8	AK073747	Oryza sat
42	94	3.8	597	12	AY195855	Synthetic
43	92.6	3.7	1006	8	AK060220	Oryza sat
44	92.6	3.7	2251	8	AK103704	Oryza sat
45	91.4	3.6	3323	6	AX746788	Sequence

ALIGNMENTS

RESULT 1
BD269560
LOCUS BD269560 2505 bp DNA linear PAT 17-JUL-2003
DEFINITION Novel CAR protein and nucleic acid molecules and uses therefor.
ACCESSION BD269560
VERSION BD269560.1 GI:33079328
KEYWORDS JP 2002535962-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2505)
AUTHORS Raju, J.
TITLE Novel CAR protein and nucleic acid molecules and uses therefor
JOURNAL Patent: JP 2002535962-A 2 29-OCT-2002;

Pred. No. is the number of results predicted by chance to have a

Db 1741 ACACGCCAATTATACATCGTGACITGACAGTCACAATATTTCTCTATGAGATGGG 1800
Qy 1801 CATGCTGTGGTGCGAGATTTGGAGAAATCAAGATTTCTACAGTCTCTCGATGAGACAAC 1860
Db 1801 CATGCTGTGGTGCGAGATTTGGAGAAATCAAGATTTCTACAGTCTCTCGATGAGACAAC 1860
Qy 1861 ATGACAAAACCAACCTGGGAACCTCCGTTGGATGGCTCCTGAGGTGTTCCAGCGAGTGCACT 1920
Db 1861 ATGACAAAACCAACCTGGGAACCTCCGTTGGATGGCTCCTGAGGTGTTCCAGCGAGTGCACT 1920
Qy 1921 CGGTACACCAATCAAAACGAGATCTTTCAGCTATGCTGTGTGTTCTGTGGAAATTTCTCACT 1980
Db 1921 CGGTACACCAATCAAAACGAGATCTTTCAGCTATGCTGTGTGTTCTGTGGAAATTTCTCACT 1980
Qy 1981 GGGGAATTTCCATTCGCTCATCTCAAGCCAGCGGCTGGGCGACGAGATGGCTTACCAC 2040
Db 1981 GGGGAATTTCCATTCGCTCATCTCAAGCCAGCGGCTGGGCGACGAGATGGCTTACCAC 2040
Qy 2041 CACATCAGACCTCCCATTTGGCTATTCATTTCCAAAGCCCATATCATCTCTGCTGATCGA 2100
Db 2041 CACATCAGACCTCCCATTTGGCTATTCATTTCCAAAGCCCATATCATCTCTGCTGATCGA 2100
Qy 2101 GGGTGAAGCGCATGTCTGAAGAGACCCGAATTTCTGAAGTTGTCTATGAAGTTAGAA 2160
Db 2101 GGGTGAAGCGCATGTCTGAAGAGACCCGAATTTCTGAAGTTGTCTATGAAGTTAGAA 2160
Qy 2161 GAGTGTCTCTGCAACATTGAGTGATGTCTCTGATCAATCAAGTCAAGCAGTGGGTCTCTC 2220
Db 2161 GAGTGTCTCTGCAACATTGAGTGATGTCTCTGATCAATCAAGTCAAGCAGTGGGTCTCTC 2220
Qy 2221 TCACCTTCT 2280
Db 2221 TCACCTTCT 2280
Qy 2281 GCAGCATTAAGAGTGGTTTCCGAATTTGGAATATGCTCTAAGTCAAGTCTCTCTCTCTCTCT 2340
Db 2281 GCAGCATTAAGAGTGGTTTCCGAATTTGGAATATGCTCTAAGTCAAGTCTCTCTCTCTCTCT 2340
Qy 2341 TTGTCCTCCAAAGTCTGGACAATTTCTCTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
Db 2341 TTGTCCTCCAAAGTCTGGACAATTTCTCTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
Qy 2401 AGTCTTCAATACACCCATTGACAAATATGCTATGATATCGATCCCATGAGTCTCATG 2460
Db 2401 AGTCTTCAATACACCCATTGACAAATATGCTATGATATCGATCCCATGAGTCTCATG 2460
Qy 2461 CATTTTCATTCTTCCGGAATAGTAGCAGCTTTTGAGGACAGCAGC 2505
Db 2461 CATTTTCATTCTTCCGGAATAGTAGCAGCTTTTGAGGACAGCAGC 2505

RESULT 2
AR269112
LOCUS AR269112 2505 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 3 from patent US 6500654.
ACCESSION AR269112
VERSION AR269112.1 GI:29699955
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2505)
AUTHORS Raju,J.
TITLE CARP protein and nucleic acid molecules and uses therefor
JOURNAL Patent: US 6500654-A 3 31-DEC-2002;
FEATURES Location/Qualifiers
source 1..2505
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 2505; DB 6; Length 2505;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGAAATTTAAATCTAGACCAACCCAAACTTTCTACTGATGAATGGAAGAAAAGTC 60
Db 1 ATGGGAAATTTAAATCTAGACCAACCCAAACTTTCTACTGATGAATGGAAGAAAAGTC 60
Qy 61 AGTGAATCATATGTTTATCAATAGAAAGATTAGAAAGTACCTGCGATCAAGGAAAAA 120
Db 61 AGTGAATCATATGTTTATCAATAGAAAGATTAGAAAGTACCTGCGATCAAGGAAAAA 120
Qy 121 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180
Db 121 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180
Qy 181 AATTACCGCAGCTGAAATGCGGCTGCTCTACTCTATTTATGTCATTTGTCATTTGTCAGGCAAG 240
Db 181 AATTACCGCAGCTGAAATGCGGCTGCTCTACTCTATTTATGTCATTTGTCATTTGTCAGGCAAG 240
Qy 241 AAATCACATATTTGAAAGGGCTCCGCCCATCTCGACTGACAAGAAAT 300
Db 241 AAATCACATATTTGAAAGGGCTCCGCCCATCTCGACTGACAAGAAAT 300
Qy 301 GGAATTTACAGCTTGGATTTAGCAGTTTACAAGATTAATGAGAAATGATCATTCTCTG 360
Db 301 GGAATTTACAGCTTGGATTTAGCAGTTTACAAGATTAATGAGAAATGATCATTCTCTG 360
Qy 361 CTTTCAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTCCTCCATATT 420
Db 361 CTTTCAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTCCTCCATATT 420
Qy 421 GCTACAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGAGCTAAATGTC 480
Db 421 GCTACAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGAGAGCTAAATGTC 480
Qy 481 AATATTTCAAGATGCAAGTTTCTTCTCACTCCATTGCAATTTGCAAGCTATGGAATGAA 540
Db 481 AATATTTCAAGATGCAAGTTTCTTCTCACTCCATTGCAATTTGCAAGCTATGGAATGAA 540
Qy 541 CAGGTAACCTCGCTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGGA 600
Db 541 CAGGTAACCTCGCTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGGA 600
Qy 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTGAATATTTGCAAACTCTTGATG 660
Db 601 GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTGAATATTTGCAAACTCTTGATG 660
Qy 661 GAAGAGGCGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCTCCATCCAT 720
Db 661 GAAGAGGCGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCTCCATCCAT 720
Qy 721 TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTGAAA 780
Db 721 TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTGAATTTGAAA 780
Qy 781 GTTCAACCTCATGTTGTTAAATATCTATGAGATACCCCTTTACACCTGCGATGCTACAAT 840
Db 781 GTTCAACCTCATGTTGTTAAATATCTATGAGATACCCCTTTACACCTGCGATGCTACAAT 840
Qy 841 GGCATAATTTGAAAGTTGCCAAGGAAATCATCAAAATATCAGGAACAGAAAGTCTGACTAAG 900
Db 841 GGCATAATTTGAAAGTTGCCAAGGAAATCATCAAAATATCAGGAACAGAAAGTCTGACTAAG 900
Qy 901 GAAACATCTTTCAGTGAACACAGCTTTTCATAGTGTCTGCTATGTCGCAAGCATTTGAC 960
Db 901 GAAACATCTTTCAGTGAACACAGCTTTTCATAGTGTCTGCTATGTCGCAAGCATTTGAC 960
Qy 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCATAAATCATCAACCAAGGAGGATGGG 1020
Db 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCATAAATCATCAACCAAGGAGGATGGG 1020
Qy 1021 CACACTGGATTAACCTCTGCTGCTGCTACCAAGTCAATTCGCTGGTTCACTTCTTACTG 1080
Db 1021 CACACTGGATTAACCTCTGCTGCTGCTACCAAGTCAATTCGCTGGTTCACTTCTTACTG 1080

Db 1021 CACACTGGATTACACTCTGCTTGTCTACACCGGTCAATTGCGCTTCAGTTCTTACTG 1080
Qy 1081 GATTAATGGAGCTGATATGAATCTAGTGGCTTTGTGTGATCCAGCAGGCTCTAGTGGTGAATA 1140
Db 1081 GATTAATGGAGCTGATATGAATCTAGTGGCTTTGTGTGATCCAGCAGGCTCTAGTGGTGAATA 1140
Qy 1141 GATCAGCAGACATGTTTGTGTGGCTTATGAAAAAGGCGATGATCCCAATGTCACACTC 1200
Db 1141 GATCAGCAGACATGTTTGTGTGGCTTATGAAAAAGGCGATGATCCCAATGTCACACTC 1200
Qy 1201 CTGAAGCATTTAAGAGACCAACAAGATGAATGCCCTGTAAATGAATATTCTCAGCTGGA 1260
Db 1201 CTGAAGCATTTAAGAGACCAACAAGATGAATGCCCTGTAAATGAATATTCTCAGCTGGA 1260
Qy 1261 GGAGATGGCTCTATGTGTCTGTTCCATCACCTTGGGGAAGATTAAAGCATGACAAA 1320
Db 1261 GGAGATGGCTCTATGTGTCTGTTCCATCACCTTGGGGAAGATTAAAGCATGACAAA 1320
Qy 1321 GAGAAGGCAGATATTCTCCTCCTAAGAGCTGGAATGCCCTTCAATTTCCATCTTCAGCTC 1380
Db 1321 GAGAAGGCAGATATTCTCCTCCTAAGAGCTGGAATGCCCTTCAATTTCCATCTTCAGCTC 1380
Qy 1381 TCAGAAATGAGTTCCATGAGATATTGGCTCAGGTTCTTTTGGGAAGTATATAAGA 1440
Db 1381 TCAGAAATGAGTTCCATGAGATATTGGCTCAGGTTCTTTTGGGAAGTATATAAGA 1440
Qy 1441 CGATGCGAATAAATAAGTGGCTATATAAAGCTTATCGAGCCAAATACCTACTGCTCCAAG 1500
Db 1441 CGATGCGAATAAATAAGTGGCTATATAAAGCTTATCGAGCCAAATACCTACTGCTCCAAG 1500
Qy 1501 TCAGATGTGATGTTTTCGCGAGAGGTGTCAATCTCTGCCAGCTCAATCATCCCTGC 1560
Db 1501 TCAGATGTGATGTTTTCGCGAGAGGTGTCAATCTCTGCCAGCTCAATCATCCCTGC 1560
Qy 1561 GTAATTCAGTTTGGGTGCTTCTGTAATGATCCAGCCAGTTGCCAATTTGCACTCA 1620
Db 1561 GTAATTCAGTTTGGGTGCTTCTGTAATGATCCAGCCAGTTGCCAATTTGCACTCA 1620
Qy 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCCTTCATGAGCAGAAGAGATTCTGATTTG 1680
Db 1621 TACATATCAGGGGTTCTCTGTTCTCCCTCCTTCATGAGCAGAAGAGATTCTGATTTG 1680
Qy 1681 CAGTCTAAATTAATTAATGAGTAGATGTTGCCAAGGCAATGAGTACCTTCAACCTG 1740
Db 1681 CAGTCTAAATTAATTAATGAGTAGATGTTGCCAAGGCAATGAGTACCTTCAACCTG 1740
Qy 1741 ACAGAGCAATATACATCGTACTTGAACAGTCAACAATATTCTCTATGAGGATGG 1800
Db 1741 ACAGAGCAATATACATCGTACTTGAACAGTCAACAATATTCTCTATGAGGATGG 1800
Qy 1801 CATGCTGTGGGAGATTTTGGAGAAATCAAGATTCTACAGTCTCTGATGAAGACAAC 1860
Db 1801 CATGCTGTGGGAGATTTTGGAGAAATCAAGATTCTACAGTCTCTGATGAAGACAAC 1860
Qy 1861 ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCCTGAGGTGTTTCAAGAGTGCAC 1920
Db 1861 ATGACAAAACAACTGGGAACCTCCGTTGGATGGCTCCTGAGGTGTTTCAAGAGTGCAC 1920
Qy 1921 CGGTACACCAATCAAGAGATGTTTTCAGCTATGCTCTGTGTGTTGGGAAATTTCTCACT 1980
Db 1921 CGGTACACCAATCAAGAGATGTTTTCAGCTATGCTCTGTGTGTTGGGAAATTTCTCACT 1980
Qy 1981 GGCGAAATTCATTCGCTCATCTCAAGCAGCGCTGCGCAGCAGACATGCTTACCAC 2040
Db 1981 GGCGAAATTCATTCGCTCATCTCAAGCAGCGCTGCGCAGCAGACATGCTTACCAC 2040
Qy 2041 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCAATATCATCTCTGTGATAGA 2100
Db 2041 CACATCAGACCTCCCATTTGGCTATTCCATTTCCCAAGCCCAATATCATCTCTGTGATAGA 2100
Qy 2101 GGGTGGAAACGATGCTCTGAGGAAGACCGGAATTTTCTGAAGTCTCATGAGTTAGAA 2160

Qy 2161 GAGTGTCTTGCAACATTGAGCTGATGTCTCTGATCAAGTAACAGCAGTGGGTCTCTC 2220
Db 2161 GAGTGTCTTGCAACATTGAGCTGATGTCTCTGATCAAGTAACAGCAGTGGGTCTCTC 2220
Qy 2221 TCACCTTCTTCTTCTTCTGATTCCCTGCTGGAACCGGGGAGGACCTGGCCGAGTCATGTG 2280
Db 2221 TCACCTTCTTCTTCTTCTGATTCCCTGCTGGAACCGGGGAGGACCTGGCCGAGTCATGTG 2280
Qy 2281 GCAGCATTAAGAGTGGTTTTCGAATTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2340
Db 2281 GCAGCATTAAGAGTGGTTTTCGAATTGGAATATGCTCTAAATGCAAGTCTCTATGCTGCT 2340
Qy 2341 TTGTCCCAAGTGTGCGACAATATTCTCTCAAGGTCTGTCTTTGGAGGAGATGAAAAGA 2400
Db 2341 TTGTCCCAAGTGTGCGACAATATTCTCTCAAGGTCTGTCTTTGGAGGAGATGAAAAGA 2400
Qy 2401 AGTCTTCAATACACCCATTGACAAATATGCTATGATCCGATCCCATGAGTCAATG 2460
Db 2401 AGTCTTCAATACACCCATTGACAAATATGCTATGATCCGATCCCATGAGTCAATG 2460
Qy 2461 CATTTTCATTTGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505
Db 2461 CATTTTCATTTGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505

RESULT 3
AR437546
LOCUS AR437546 2505 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 3 from patent US 6660490.
ACCESSION AR437546
VERSION AR437546.1 GI:40202641
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2505)
AUTHORS Raju,J.
TITLE CARB protein and nucleic acid molecules and uses therefor
JOURNAL Patent: US 6660490-A 3 09-DEC-2003;
FEATURES
source Location/Qualifiers
1..2505
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 2505; DB 6; Length 2505;
Best local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGGAAATTAATAATCTAGACCAACCCAAACTTGTACTGTAGTGAATGGAAGAAAAAGTC 60
Db 1 ATCGGAAATTAATAATCTAGACCAACCCAAACTTGTACTGTAGTGAATGGAAGAAAAAGTC 60
Qy 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGAAAAA 120
Db 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGAAAAA 120
Qy 121 GAATGACAGACTAGGAATATTTGCTCTGATGAAGCTTCAGTAAAGTCAATTTA 180
Db 121 GAATGACAGACTAGGAATATTTGCTCTGATGAAGCTTCAGTAAAGTCAATTTA 180
Qy 181 AATTACCGCACTGAAAATGGGCTGTCTTACTTCTATTATGTTGCAATTTGTGGAGGCAAG 240
Db 181 AATTACCGCACTGAAAATGGGCTGTCTTACTTCTATTATGTTGCAATTTGTGGAGGCAAG 240
Qy 241 AATATCATATTCGAACTCTTATGTTGAAGGGCTCCGCCATCTCGACTGACAGAAAT 300
Db 241 AATATCATATTCGAACTCTTATGTTGAAGGGCTCCGCCATCTCGACTGACAGAAAT 300
Qy 301 GGATTTACAGCCTTGTGATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG 360
Db 301 GGATTTACAGCCTTGTGATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG 360

RESULT 4
AX056437
LOCUS AX056437 2508 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 81 from Patent WO0073469.
ACCESSION AX056437
VERSION AX056437.1 GI:12229144
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plozman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 81 07-DEC-2000;
Sugen, Inc. (US)
FEATURES
source
Location/Qualifiers
1..2508
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 2505; DB 6; Length 2508;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGAATTATTAATCTAGACCAACCCAACTTGACTGATGAATGGAAGAAAAGTC 60
DB |||||
QY 1 ATGGGAATTATTAATCTAGACCAACCCAACTTGACTGATGAATGGAAGAAAAGTC 60
DB |||||
QY 61 AGTGAATCATATGTTATCACATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 120
DB |||||
QY 61 AGTGAATCATATGTTATCACATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 120
DB |||||
QY 121 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180
DB |||||
QY 121 GAATGACAGAACTAAGAAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180
DB |||||
QY 181 AATTACCGCACTGAAATGGGCTCTCTACTTATTTGATGATGTTGAGGCAAG 240
DB |||||
QY 181 AATTACCGCACTGAAATGGGCTCTCTACTTATTTGATGATGTTGAGGCAAG 240
DB |||||
QY 241 AAATCACATATTCGAATCTTTATGTTGAAAGGGCTCGGCCATCTCGACTGCAAGAAAT 300
DB |||||
QY 241 AAATCACATATTCGAATCTTTATGTTGAAAGGGCTCGGCCATCTCGACTGCAAGAAAT 300
DB |||||
QY 301 GGATTTACAGCTTTGATTTAGCAGTTTACAAGGATATGCAAGTGTGATGATCTCTG 360
DB |||||
QY 301 GGATTTACAGCTTTGATTTAGCAGTTTACAAGGATATGCAAGTGTGATGATCTCTG 360
DB |||||
QY 361 CTTACAGTGGAGCTGATATACAGCAGTTGGATAGCGTGGCTCACTGCCCTCCATATT 420
DB |||||
QY 361 CTTACAGTGGAGCTGATATACAGCAGTTGGATAGCGTGGCTCACTGCCCTCCATATT 420
DB |||||
QY 421 GCTCAATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGGAGCTAATGTC 480
DB |||||
QY 421 GCTCAATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATGGAGCTAATGTC 480
DB |||||
QY 481 AATATTCAAGATGAGTGTATTTTACCTCCATTTGATGATGATGCAAGCTGATGATGAA 540
DB |||||
QY 481 AATATTCAAGATGAGTGTATTTTACCTCCATTTGATGATGATGCAAGCTGATGATGAA 540
DB |||||
QY 541 CAGGTAACTCGCTTTCTTTGAAATTTGGTGTGATGATGATGATGATGATGATGATGATG 600
DB |||||
QY 541 CAGGTAACTCGCTTTCTTTGAAATTTGGTGTGATGATGATGATGATGATGATGATGATG 600
DB |||||
QY 601 GATAGACCCCTCCACCTAGATGCTGCAAAAGGATTTCTGAATATTGCAAACTCTTGATG 660
DB |||||
QY 601 GATAGACCCCTCCACCTAGATGCTGCAAAAGGATTTCTGAATATTGCAAACTCTTGATG 660
DB |||||
QY 661 GAAGAGGCGCAAGAGAGATGTAATGCTCAAGATAATGAAGACCATGTCCTCCATCCAT 720
DB |||||

DB 661 GAAGAGGCGCAAGAGAGATGTAATGCTCAAGATAATGAAGACCATGTCCTCCATCCAT 720
QY TTCTGTTCTCGATTGGACACCATGATATAGTTAACTATCTGCTGCAAAAGTGAATTTGAA 780
DB |||||
QY TTCTGTTCTCGATTGGACACCATGATATAGTTAACTATCTGCTGCAAAAGTGAATTTGAA 780
DB |||||
QY GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGCACTGCTACCAAT 840
DB |||||
QY GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACCTGCACTGCTACCAAT 840
DB |||||
QY GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
DB |||||
QY GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
DB |||||
QY GAAACATCTTTCAGTGAACAGCTTTTCATAGTGTGTTACCTATGCGAAGAGCATTTGAC 960
DB |||||
QY GAAACATCTTTCAGTGAACAGCTTTTCATAGTGTGTTACCTATGCGAAGAGCATTTGAC 960
DB |||||
QY CTAGTCAAAATTTCTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB |||||
QY CTAGTCAAAATTTCTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB |||||
QY CACATGGAATTACACTCTGCTTGTCTACCAAGTCAATTCGCTGCTGCTGCTGCTGCTGCTG 1080
DB |||||
QY CACATGGAATTACACTCTGCTTGTCTACCAAGTCAATTCGCTGCTGCTGCTGCTGCTGCTG 1080
DB |||||
QY GATAATGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
DB |||||
QY GATAATGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
DB |||||
QY GATGAGCAGACATGTTTGTGTTGCTTATGAAAGGGCATGATGATGATGATGATGATGATG 1200
DB |||||
QY GATGAGCAGACATGTTTGTGTTGCTTATGAAAGGGCATGATGATGATGATGATGATGATG 1200
DB |||||
QY CTGAAGCATTAAGAGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
DB |||||
QY CTGAAGCATTAAGAGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
DB |||||
QY GGAGATGGCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB |||||
QY GGAGATGGCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB |||||
QY GAGAAGGCGAGATATCTCTCTTAAGAGCTGATGATGATGATGATGATGATGATGATGATG 1380
DB |||||
QY GAGAAGGCGAGATATCTCTCTTAAGAGCTGATGATGATGATGATGATGATGATGATGATG 1380
DB |||||
QY TCAGAAATTTGAGTTCCATGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB |||||
QY TCAGAAATTTGAGTTCCATGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB |||||
QY CGATGAGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTACTGCTCCAAAG 1500
DB |||||
QY CGATGAGCAAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTACTGCTCCAAAG 1500
DB |||||
QY TCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
DB |||||
QY TCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
DB |||||
QY GTAAATTCAGTTTGTGGTCTTGTGTTGATGATGATGATGATGATGATGATGATGATGATG 1620
DB |||||
QY GTAAATTCAGTTTGTGGTCTTGTGTTGATGATGATGATGATGATGATGATGATGATGATG 1620
DB |||||
QY TCATATCAGGGGGTCTCTGTTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
DB |||||
QY TCATATCAGGGGGTCTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
DB |||||
QY CAGTCTAAATTAATTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB |||||
QY CAGTCTAAATTAATTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB |||||
QY ACACAGCCAAATTAATCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB |||||
QY ACACAGCCAAATTAATCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB |||||

```
QY 1801 CATGCTGTGGTGGCAGATTTTGGAGAACTCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
Db 1801 CATGCTGTGGTGGCAGATTTTGGAGAACTCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
QY 1861 ATGACAAAACAACCTGGGAACCTCCGTTGGATGGCTCTCAGGTTTCCACAGTGCAC 1920
Db 1861 ATGACAAAACAACCTGGGAACCTCCGTTGGATGGCTCTCAGGTTTCCACAGTGCAC 1920
QY 1921 CGGTACACCATCAAAACAGATGCTTTCAGTATGCTCTGTCTGTCTGTGGAAATTCCTACT 1980
Db 1921 CGGTACACCATCAAAACAGATGCTTTCAGTATGCTCTGTCTGTCTGTGGAAATTCCTACT 1980
QY 1981 GGGGAAATTCATTCCTCTCATCTCAAGCAGCGGCTGGCGCAGCAGACATGGCTTACCAAC 2040
Db 1981 GGGGAAATTCATTCCTCTCATCTCAAGCAGCGGCTGGCGCAGCAGACATGGCTTACCAAC 2040
QY 2041 CATACAGACCTCCCATTTGCTATTCCTCAAGCAGCGGCTGGCGCAGCAGACATGGCTTACCAAC 2100
Db 2041 CATACAGACCTCCCATTTGCTATTCCTCAAGCAGCGGCTGGCGCAGCAGACATGGCTTACCAAC 2100
QY 2101 GGGTGAACGCGATGCTCAAGGAGACCCGGAATTTCTGAAGTTGTCATGAAGTTAGAA 2160
Db 2101 GGGTGAACGCGATGCTCAAGGAGACCCGGAATTTCTGAAGTTGTCATGAAGTTAGAA 2160
QY 2161 GAGTGTCTCTGCAACATGAGCTGATGCTCTCTGCAATCAAGTAACAGCAGTGGTCTCTC 2220
Db 2161 GAGTGTCTCTGCAACATGAGCTGATGCTCTCTGCAATCAAGTAACAGCAGTGGTCTCTC 2220
QY 2221 TCACCTCTCTCTCTCTCTGATGCTGCTGGAACCCGGGAGGACCTGGCGGAGTCACTG 2280
Db 2221 TCACCTCTCTCTCTCTGATGCTGCTGGAACCCGGGAGGACCTGGCGGAGTCACTG 2280
QY 2281 GCAGCATTAAGAAAGTGGTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTCT 2340
Db 2281 GCAGCATTAAGAAAGTGGTTCGAATTTGGAATATGCTCTAAATGCAAGTCTCTATGCTCT 2340
QY 2341 TTGTCCCAAGTCTGGCAATATTCCTCTCAAGTCTGCTTTGGAGAGATGAAGA 2400
Db 2341 TTGTCCCAAGTCTGGCAATATTCCTCTCAAGTCTGCTTTGGAGAGATGAAGA 2400
QY 2401 AGTCTTCAATACACACCATGACAAATATGCTATGCTATGCTATGCTATGCTATGCTATG 2460
Db 2401 AGTCTTCAATACACACCATGACAAATATGCTATGCTATGCTATGCTATGCTATGCTATG 2460
QY 2461 CATTTTCATTTCTGGCGAAATAGTACAGCTTTGAGGACAGCAGC 2505
Db 2461 CATTTTCATTTCTGGCGAAATAGTACAGCTTTGAGGACAGCAGC 2505

RESULT 5
AY303691 2980 bp mRNA linear PRI 23-JUN-2003
LOCUS Homo sapiens cardiac ankyrin repeat kinase mRNA, complete cds.
DEFINITION
ACCESSION AV303691
VERSION AV303691.1 GI:32165605
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2980)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Jeyaseelan,R.
JOURNAL Cardiac Ankyrin Repeat Kinase (CARK)
REFERENCE 2 (bases 1 to 2980)
AUTHORS Unpublished
TITLE Jeyaseelan,R.
JOURNAL Direct Submission
TITLE Submitted (20-MAY-2003) Cardiovascular Biology, Millennium
JOURNAL Pharmaceuticals Inc., 75 Sidney Street, Cambridge, MA 02139, USA
FEATURES
source Location/Qualifiers
1. .2980
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p22.3-p13.3"
31.2538
/note="CARK; protein kinase"
/codon_start=1
/product="cardiac ankyrin repeat kinase"
/protein_id="AAP72030.1"
/db_xref="GI:32165606"
/translation="MGNYKSRPTQCTDDEWKKVSESVYITIERLEDDLOIKEXELTE
LENIFGSEARSKNLYRFTENGSLHLCCICGKSKSHRTLMKGLRPSRLTRNGP
TALHAYTKDNABLITSLHSGADIQQVGYGGLTALHATAGHLEADVLLQHGAVV
NIDAVFTTFLHIAAYGHEQVTRLLKFGADVNVGEGDRPLHLASAKGFLNIAKL
LMEEKSKADVNAQDNEDHVLHFCRFGHDIKYLQSLDEVPVHVNYIGDTPHLH
ACYNGKFEVAKELIIQISGTESTXENIFSETAFHSACTYKSIDLVKFLDQNVIN
HGRDHTGLHSACVGHIRLVQFLDNGADMNIVACDPSRSGEKDEQTCIMWAYEK
GHDAIVTLKXKVRQDELPCNEYSGGDCGYSVPSPLGKIKSMTKKADILLRA
GLPSHFHLQLEIEFERHIIIGSGFGKVKYKGRNKIIVALKYRANTYKCSKSDVDFCR
EYVILCQUNHPCVIOFGACLNDFSPFAITVQYISGSLFSLHLEHQRLDLOSKLII
AVDVAKGMEYHLNLTQPIIHRDLNSHNILLYEDGHVAVDEGSRFLQSLDQNTKQ
PGLNRMAPVFTTQCTRTYIKADVFSYALCUIWEITGEIPFAHLKPAADAAWHHI
RNPIGYSIPKPISSILLIRGMNACPEGRPEFSEVVMKLEECLELMSPPASSNSGSL
SPSSSDCLVNRGGPGRSHVAALRSRPELEYALNARSVAALSQSGQVSSQGLSLEBM
KRSLOYTPIDKYGYYVSDPMSMHPHSCNRSNSSFEDSS"
```

ORIGIN

QY	1	ATGGGAAATTAATAATCTAGACCAACCCAACTTGTCTAGTGAATGAAGAAAAAGTC	60
Db	31	ATGGGAAATTAATAATCTAGACCAACCCAACTTGTCTAGTGAATGAAGAAAAAGTC	90
QY	61	AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGATGACCTGCAGATCAAGAAAAA	120
Db	91	AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGATGACCTGCAGATCAAGAAAAA	150
QY	121	GAATCAGAGAACTAAGGAATATATTTGGCTCTGATGAAGCTTTCAGTAAAGTCAATTTA	180
Db	151	GAATCAGAGAACTAAGGAATATATTTGGCTCTGATGAAGCTTTCAGTAAAGTCAATTTA	210
QY	181	AATTACCGCACTGAAATGGGCTGCTCTACTTATTTATGTTGATTTGTTGGAGGCAAG	240
Db	211	AATTACCGCACTGAAATGGGCTGCTCTACTTATTTATGTTGATTTGTTGGAGGCAAG	270
QY	241	AAATCAATATTGCAACTCTTATGTTGAAAGGCTCCGCCATCTCGACTGACAGAAAT	300
Db	271	AAATCAATATTGCAACTCTTATGTTGAAAGGCTCCGCCATCTCGACTGACAGAAAT	330
QY	301	GGATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCAGAAATTCATCTCTCTG	360
Db	331	GGATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCAGAAATTCATCTCTCTG	390
QY	361	CTTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTCACTGCCCTCATATT	420
Db	391	CTTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCCTCACTGCCCTCATATT	450
QY	421	GCTACAAATAGTGGCCACCTAGAGGCTGCTGATGCTGCTGTTGCAACATGGAGCTAATGTC	480
Db	451	GCTACAAATAGTGGCCACCTAGAGGCTGCTGATGCTGCTGTTGCAACATGGAGCTAATGTC	510
QY	481	AATATTCAAGATGCAAGTTTTTTTCTACTCCATTTGCATATTTGCAGCGTACTATGGACATGAA	540
Db	511	AATATTCAAGATGCAAGTTTTTTTCTACTCCATTTGCATATTTGCAGCGTACTATGGACATGAA	570
QY	541	CAGGTAACCTCGCTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA	600
Db	571	CAGGTAACCTCGCTCTTTTGAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA	630
QY	601	GATAGACCCCTCCACCTAGCATCTGCAGAAAGGATTTCTTGAATATTGCAAAATCTTGTATG	660

source

```
1. .3001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HH498"
/tissue_type="heart"
/dev_stage="adult"
52. -2559
/codon_start=1
/product="putative protein-tyrosine kinase"
/protein_id="A428632.1"
/db_xref="GI:4768829"
```

CDS

```
/translation="MGNYKSRPTCTDEWKKKVSSEVIVIERLEDLQIKEKELTE
LNIFGSDPAFKNLVNRTENGLSLHLCICCGKKSHIRTLMLKLRPSRLTRNGF
TALHLAVYKDNALSLTSLHSGADIQVYGGVGLTALTIAGHLEADSVLLQHGAV
N/QDAVFFPLHIAAYGHEQVTRLLKFGADVNVSGEVRPLHLASAKGFLNIAKL
LMEBSKADVNQDNEDHVP/LHFCRFCHHDI/VKYLLOSLDEVOHPVNIYIGDTP/LHL
ACTNGKFEVAKETIIQISGESLTKENISETAFSACTYGSISILVKFLIDONVININ
HGRDGHGT/LHSACTHGHIRLVQFLDNGADNVLVADPSRSGSEKBEQICLWAYEK
GHDAIVTLKHYRFEIPEIENYQPGDGSVYVSP/LGKIKMSIKTEKADILL/LRA
GLPSFHLQSLSEIEIIGSGDFGKVKGRNKAIVAKRYRANTYCSKSDVDMFR
EVSILCOLNHPCVIOFVGAICLNDPSQFAIVTQYISGSLFSLHLEQRILDLQSLKI
AVDVAKMEVYLNLCPIIHRDLNLSNILLVEDGHAVVADFGESRFLQSLDEDNMTKO
PGLNRMAEVTQCTRYTIKADVSVALCLWEILTGEI/PAHLKPAARADMAVYHI
RPPIGYSIPKISSILLIRGNWACPEGRFEFSEVNVKLECLNIELMSPASSNSGSL
SPSSSDCIVNRRGGPGRSHVAALRFRFEYALNARSYAALQSAGQISSQGLSLEM
KRSLOVTPIDKTYGVSDPMSMHFHS CRNSSSFEDES"
```

2978. .2983

polya_signal
ORIGIN

Query Match 100.0%; Score 2505; DB 9; Length 3001;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGAAATTATAAATCTAGACCAACCCAAACTTGCTGATGATGAATGGAAGAAAAGTC 60
DB 52 ATGGGAAATTATAAATCTAGACCAACCCAAACTTGCTGATGATGAATGGAAGAAAAGTC 111
QY 61 AGTGAATCATATGTTATCAATAGAAAAGATTAGAAAGATGACCTGCAGATCAAGAAAAA 120
DB 112 AGTGAATCATATGTTATCAATAGAAAAGATTAGAAAGATGACCTGCAGATCAAGAAAAA 171
QY 121 GAATGACGAACTAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180
DB 172 GAATGACGAACTAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 231
QY 181 AATTACGCACTGAAAATGGGCTGTCTCACTTCAATTTATGTTGCAATTTGTGGAGGCAAG 240
DB 232 AATTACGCACTGAAAATGGGCTGTCTCACTTCAATTTATGTTGCAATTTGTGGAGGCAAG 291
QY 241 AATACATATTCGAATCTTATTTGAAAGGCTCCGCCATCTCGACTGACAAAGAAAT 300
DB 292 AATACATATTCGAATCTTATTTGAAAGGCTCCGCCATCTCGACTGACAAAGAAAT 351
QY 301 GGATTTACGCTTGCAATTTAGCAGTTTACAAGGATAATGCAGAAATTGATCACTTCTCTG 360
DB 352 GGATTTACGCTTGCAATTTAGCAGTTTACAAGGATAATGCAGAAATTGATCACTTCTCTG 411
QY 361 CTTCAAGTGGAGTGATATACAGCAGTTTGATAGCGTGGCTTCACTGCCCTCCATATT 420
DB 412 CTTCAAGTGGAGTGATATACAGCAGTTTGATAGCGTGGCTTCACTGCCCTCCATATT 471
QY 421 GCTCAATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATCGAGCTAATGTC 480
DB 472 GCTCAATAGTGGCCACCTAGAGGCTGCTGATGCTGTTGCAACATCGAGCTAATGTC 531
QY 481 AATATTCAGATGCAAGTTTTTTTCACTCCATTCGATTCGACGGTACTATGGACATGAA 540
DB 532 AATATTCAGATGCAAGTTTTTTTCACTCCATTCGATTCGACGGTACTATGGACATGAA 591
QY 541 CAGTAACTCCCTCTTTTGAATTTGGTCTGATCTAATGTAATGTAAGTGTGAAGTTGA 600
DB 592 CAGTAACTCCCTCTTTTGAATTTGGTCTGATCTAATGTAATGTAAGTGTGAAGTTGA 651
```

QY 1681 CAGTCTAAATTAATTAATGAGTGTGCGAAGGATGAGTACCTTCAACACCTG 1740
Db 1732 CAGTCTAAATTAATTAATGAGTGTGCGAAGGATGAGTACCTTCAACACCTG 1791
QY 1741 ACACAGCCAAATTAATGAGTGTGCGAAGGATGAGTACCTTCAACACCTG 1800
Db 1792 ACACAGCCAAATTAATGAGTGTGCGAAGGATGAGTACCTTCAACACCTG 1851
QY 1801 CATGCTGTGGTGGAGATTTGGAGATTCAGATTCAGATTCAGATTCAGATTC 1860
Db 1852 CATGCTGTGGTGGAGATTTGGAGATTCAGATTCAGATTCAGATTCAGATTC 1911
QY 1861 ATGACAAAACCACTGGGAACCTCGTTGGATGGCTCTGAGGTGTTCAACGAGTGCAC 1920
Db 1912 ATGACAAAACCACTGGGAACCTCGTTGGATGGCTCTGAGGTGTTCAACGAGTGCAC 1971
QY 1921 CGGTACACATCAAGAGAGATGTTTCAAGTATGCTCTGTTGTTGTTGTTGTTGTT 1980
Db 1972 CGGTACACATCAAGAGAGATGTTTCAAGTATGCTCTGTTGTTGTTGTTGTTGTT 2031
QY 1981 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGCTCGGAGAGATGTTGTTGTTGTTGTT 2040
Db 2032 GCGGAAATTCATTCGCTCATCTCAAGCCAGCGCTCGGAGAGATGTTGTTGTTGTTGTT 2091
QY 2041 CACATCAGACCTCCATTTGGTATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2100
Db 2092 CACATCAGACCTCCATTTGGTATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2151
QY 2101 GGGTGGAGAGATGTTTGGAGAGAGATTCATTCATTCATTCATTCATTCATTCATTCAT 2160
Db 2152 GGGTGGAGAGATGTTTGGAGAGAGATTCATTCATTCATTCATTCATTCATTCATTCAT 2211
QY 2161 GAGTGTCTCTCAACATTTGAGTGTCTCTCAAGTATGCTCTCAAGTATGCTCTCAAGTAT 2220
Db 2212 GAGTGTCTCTCAACATTTGAGTGTCTCTCAAGTATGCTCTCAAGTATGCTCTCAAGTAT 2271
QY 2221 TCACCT 2280
Db 2272 TCACCT 2331
QY 2281 GCAGCATTAAGAAGTCTGTTTGGAAATGGAATATGCTCTCAAGTATGCTCTCAAGTAT 2340
Db 2332 GCAGCATTAAGAAGTCTGTTTGGAAATGGAATATGCTCTCAAGTATGCTCTCAAGTAT 2391
QY 2341 TTGTCCCAAGTCTGAGATATTCCTCTCAAGTATGCTCTCAAGTATGCTCTCAAGTAT 2400
Db 2392 TTGTCCCAAGTCTGAGATATTCCTCTCAAGTATGCTCTCAAGTATGCTCTCAAGTAT 2451
QY 2401 AGTCTTCAATACACACCATTTGAAATATGGAATATGGAATATGGAATATGGAATAT 2460
Db 2452 AGTCTTCAATACACACCATTTGAAATATGGAATATGGAATATGGAATATGGAATAT 2511
QY 2461 CATTTTCATTTCCCGAAATAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2505
Db 2512 CATTTTCATTTCCCGAAATAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2556

RESULT 7

BD269559 3025 bp DNA linear PAR 17-JUL-2003
DEFINITION Novel CARP protein and nucleic acid molecules and uses thereof.
ACCESSION BD269559
VERSION BD269559.1 GI:33079327
KEYWORDS JP 2002535962-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Raju, J.
TITLE Novel CARP protein and nucleic acid molecules and uses thereof
JOURNAL Patent: JP 2002535962-A 1 29-OCT-2002;
MILLENNIUM PHARMACEUTICALS INC

COMMENT

OS Homo sapiens (human)
PN JP 2002535962-A/1
PD 29-OCT-2002
PF 10-DEC-1999 JP 2000586772
PR 11-DEC-1998 US 60/111938,14-APR-1999 US 09/291839 P1
JEYASBELAN RAJU
PC C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N9/10,C1201/02,
PC C12Q1/68,G01N33/15,G01N33/50,G01N33/53,G01N33/68//A61K38/45,
PC A61K39/395,
PC A61K45/00,A61P9/04,A61P9/10,A61P35/00,A61P43/00,C12N15/00, PC
C12N5/00,
PC A61K37/52
CC Novel CARP protein and nucleic acid molecules and uses CC
therefor
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..3025
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES

source
ORIGIN

Query Match 100.0%; Score 2505; DB 6; Length 3025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGAATTAATTAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAATGTC 60
Db 48 ATGGGAATTAATTAATCTAGACCAACCCAACTTGTACTGATGAATGGAAGAAAATGTC 107
QY 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGATGACCTCGATCAAGGAAAA 120
Db 108 AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGATGACCTCGATCAAGGAAAA 167
QY 121 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTA 180
Db 168 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCCTCAGTAAAGTCAATTA 227
QY 181 AATTACCGCACTGAAATGGGCTCTCTACCTTCAATTTATGTTGATTTGTTGGAGCAAG 240
Db 228 AATTACCGCACTGAAATGGGCTCTCTACCTTCAATTTATGTTGATTTGTTGGAGCAAG 287
QY 241 AATTACCAATTCGAATCTTATGTTGAAAGGGCTCCGCCCACATCTCGACTGACAGAAAT 300
Db 288 AATTACCAATTCGAATCTTATGTTGAAAGGGCTCCGCCCACATCTCGACTGACAGAAAT 347
QY 301 GGATTTACAGCCCTTGCATTTAGCAGTTTACAGGATTAATGCAAGATTCATCCTCTCTG 360
Db 348 GGATTTACAGCCCTTGCATTTAGCAGTTTACAGGATTAATGCAAGATTCATCCTCTCTG 407
QY 361 CTTTACAGTGGAGCTGATATACAGCAGTTGATGATCGGTGGCTCCTCAGTCCCTCCATTT 420
Db 408 CTTTACAGTGGAGCTGATATACAGCAGTTGATGATCGGTGGCTCCTCAGTCCCTCCATTT 467
QY 421 GCTACAAATAGTGGCCACCTAGAGGCTGCTGATGCTGCTTGGCAACATGGAGCTAATGTC 480
Db 468 GCTACAAATAGTGGCCACCTAGAGGCTGCTGATGCTGCTTGGCAACATGGAGCTAATGTC 527
QY 481 AATATTCAAGATGCGATTTTTTTTCACTTCATTTGCAATATGCAAGGATTAATGCAAGCA 540
Db 528 AATATTCAAGATGCGATTTTTTTTCACTTCATTTGCAATATGCAAGGATTAATGCAAGCA 587
QY 541 CAGGTAATCGGCTTCTTTTGAATTTGGTGTGATGTAATGTAATGTAATGTAATGTAATG 600
Db 588 CAGGTAATCGGCTTCTTTTGAATTTGGTGTGATGTAATGTAATGTAATGTAATGTAATG 647
QY 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATGCAAAACTCTCTGATG 660
Db 648 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATGCAAAACTCTCTGATG 707
QY 661 GAAGAAGGCGCAAGAGCAGATGTGTAATGCTCAAGATAATGAAGACCATGTGCCACTCCAT 720

Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGGGAAATTTATAAATCTAGACCAACCCAAATCTGTACTGATGAATGGAAAGAAAAAGTC	60
Db	48	ATGGGAAATTTATAAATCTAGACCAACCCAAATCTGTACTGATGAATGGAAAGAAAAAGTC	107
Qy	61	AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAAA	120
Db	108	AGTGAATCATATGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAAA	167
Qy	121	GAACCTGACGAACTTAAGGAATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA	180
Db	168	GAACCTGACGAACTTAAGGAATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA	227
Qy	181	AATTACCGCTCGAATAATGGCGTCTCTACTCTATTTATGTTGCAATTTGTGGAGGCAAG	240
Db	228	AATTACCGCTCGAATAATGGCGTCTCTACTCTATTTATGTTGCAATTTGTGGAGGCAAG	287
Qy	241	AAATCACATATTCGAACCTTATGTTGAAAGGGCTCCGCCATCTCGACTCACAGAAT	300
Db	288	AAATCACATATTCGAACCTTATGTTGAAAGGGCTCCGCCATCTCGACTCACAGAAT	347
Qy	301	GGATTTACAGCTTTAGCAGTTTACAGGTTTACAAGTAATGACAGAAATGATCACTTCTCG	360
Db	348	GGATTTACAGCTTTAGCAGTTTACAGGTTTACAAGTAATGACAGAAATGATCACTTCTCG	407
Qy	361	CTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATT	420
Db	408	CTTCACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCTCACTGCCCTCCATATT	467
Qy	421	GCTACAATAGCTGGCCACCTAGAGCGTCTGATGTGCTGTTGCCAATGAGAGCTAAATGTC	480
Db	468	GCTACAATAGCTGGCCACCTAGAGCGTCTGATGTGCTGTTGCCAATGAGAGCTAAATGTC	527
Qy	481	AATATTCAAGATGCAGTTTTTTTCACTCAATTTGCATATTTGCAGGTACTATGACATGAA	540
Db	528	AATATTCAAGATGCAGTTTTTTTCACTCAATTTGCATATTTGCAGGTACTATGACATGAA	587
Qy	541	CAGTAACTCGCCTCTTTTGAATTTGGTGTGATGATGAAATGTAAGTGGTGAAGTTTGA	600
Db	588	CAGTAACTCGCCTCTTTTGAATTTGGTGTGATGATGAAATGTAAGTGGTGAAGTTTGA	647
Qy	601	GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATTTGCAAACTCTTGATG	660
Db	648	GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTCTTGAATATTTGCAAACTCTTGATG	707
Qy	661	GAAGAAGGAGCAAGACGATCTGAAATGCTCAAGATATGAAGACCATGTCACCTCCAT	720
Db	708	GAAGAAGGAGCAAGACGATCTGAAATGCTCAAGATATGAAGACCATGTCACCTCCAT	767
Qy	721	TTCTGTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATTTGGAA	780
Db	768	TTCTGTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATTTGGAA	827
Qy	781	GTTCAACCTCATGTTGTTAATCTATGGAGTACCCCTTTACACCTGGCATGCTCAAT	840
Db	828	GTTCAACCTCATGTTGTTAATCTATGGAGTACCCCTTTACACCTGGCATGCTCAAT	887
Qy	841	GGCAAAATTTGAAAGTTGCCAAGGAAATCATCCAAATATCAGGAAACAGAAAGTCTGACTAAG	900
Db	888	GGCAAAATTTGAAAGTTGCCAAGGAAATCATCCAAATATCAGGAAACAGAAAGTCTGACTAAG	947
Qy	901	GAAGACATCTTCAGTGAAGACAGCTTTTCATATGCTGTTAGTACCTATGGCAAGAGCATTTGAC	960
Db	948	GAAGACATCTTCAGTGAAGACAGCTTTTCATATGCTGTTAGTACCTATGGCAAGAGCATTTGAC	1007
Qy	961	CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAATCAACACCAAGGAGGGATGGG	1020
Db	1008	CTAGTCAAAATTTCTTCTTGATCAGAAATGTCATAAATCAACACCAAGGAGGGATGGG	1067
Qy	1021	CACACTGGATTAACCTCTGCTTGTACACCGGTCAATTCGCTGGTTCAAGTTCTTACTG	1080
Db	1068	CACACTGGATTAACCTCTGCTTGTACACCGGTCAATTCGCTGGTTCAAGTTCTTACTG	1127

QY	2161	GAGTGTCTCTGCAACATTTAGAGTCGATGTCTCTTGCAATCAAGTAACAGCAGTGGGTCTCTC	2220
Db	2208	GAGTGTCTCTGCAACATTTAGAGTCGATGTCTCTTGCAATCAAGTAACAGCAGTGGGTCTCTC	2267
QY	2221	TCACCTTCTCTTCTTCTTGATTTGCTGGTGAAACGGGGAGGACCTGGCGGAGTCATGTG	2280
Db	2268	TCACCTTCTCTTCTTCTTGATTTGCTGGTGAAACGGGGAGGACCTGGCGGAGTCATGTG	2327
QY	2281	GCAGCATTTAAGAAGTCGTTTCGAATTGGAAATATGCTCTTAAATGCAAGTCTCTATGCTGCT	2340
Db	2328	GCAGCATTTAAGAAGTCGTTTCGAATTGGAAATATGCTCTTAAATGCAAGTCTCTATGCTGCT	2387
QY	2341	TTGTCCCAAAAGTGTGGACAAATATTCTCTCTCAAGGTCTGTCTTTGGAGGAGATGAAAGA	2400
Db	2388	TTGTCCCAAAAGTGTGGACAAATATTCTCTCTCAAGGTCTGTCTTTGGAGGAGATGAAAGA	2447
QY	2401	AGTCTTCAATACACACCATTCGACAAATATGGCTATGTATCCGATCCCATGAGTCAATG	2460
Db	2448	AGTCTTCAATACACACCATTCGACAAATATGGCTATGTATCCGATCCCATGAGTCAATG	2507
QY	2461	CATTTTTCATTTCTGCCGAATATAGTAGCAGCTTTTGAGGACAGCAGC	2505
Db	2508	CATTTTTCATTTCTGCCGAATATAGTAGCAGCTTTTGAGGACAGCAGC	2552
RESULT 9			
AR437545			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			

ORIGIN

Query Match	100.0%;	Score 2505;	DB 6;	Length 3025;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2505;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGGAAATTATAACTCTAGACCAACCAACCTTGCTACTGATCAATGGAGGAAAGATC	60	
DB	48	ATGGGAAATTATAAATCTAGACCAACCAACCTTGCTACTGATCAATGGAGGAAAGATC	107	
QY	61	AGTGAATCATATGTTTATCACAAATAGAAGATTAGAAGATGACCTGCAGATCAAGGAAAAA	120	
DB	108	AGTGAATCATATGTTTATCACAAATAGAAGATTAGAAGATGACCTGCAGATCAAGGAAAAA	167	
QY	121	GAACGTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGACCTTCAGTAAAGTCAATTTA	180	
DB	168	GAACGTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGACCTTCAGTAAAGTCAATTTA	227	
QY	181	AATTACCGCACGAAATAGGGGTGCTCTCACTTCATTATGTTGGCAATTTGTGGAGGCAAG	240	
DB	228	AATTACCGCACGAAATAGGGGTGCTCTCACTTCATTATGTTGGCAATTTGTGGAGGCAAG	287	
QY	241	AAATCACATATTCGAACCTCTTATGTTGAAAGGGCTCGGCCCATCTCGACTGACAGAAAT	300	
DB	288	AAATCACATATTCGAACCTCTTATGTTGAAAGGGCTCGGCCCATCTCGACTGACAGAAAT	347	
QY	301	GGATTTACAGCCTTCGATTTAGCAGTTTACAAGATAATCGAAGTTGATCACTTCTCTG	360	
DB	348	GGATTTACAGCCTTCGATTTAGCAGTTTACAAGATAATCGAAGTTGATCACTTCTCTG	407	

[illegible]

1501	TCAGAGCTGGATATGTTTTGGCGAGAGGTGTCCTATCTCTCGCAGCTCAACCCACCCCTGC	1560
DB		
1561	GTAATTCAGTTTTGGGGTGCTTGCTTTGAATGATCCCAAGCTTTGGCATTTGCACTCAA	1620
QY		
1561	GTGGTTCAGTTTGGGGTGCTTGCTTGGAATGACCCAGTCAGTTTGGCAATTTGCACTCAG	1620
DB		
1621	TACATATCAGGGGGTCTCTGTCTTCCCTCTTTCATGAGCAGAGAGGATTTCTTGATTTG	1680
QY		
1621	TACATTTTCAGAGGCTCCCTGTTTCTCCCTGCTTCATGAACAGAGAGAAATTCITTGACTTG	1680
DB		
1681	CAGTCTAAATTAATTATTCGACGTAGATGTTCGCAAGGCATGGAGTACCTTCAACAACCTG	1740
QY		
1681	CAGTCTAAATTAATCATTCGGGTAGACGTTGCCAAGGGCATGGAGTACCTGCAACAGCTTG	1740
DB		
1741	ACACAGCCAAATTAACATCGTAGCTTGAACAGCTCACAATATTTCTTCTATGAGGATGGG	1800
QY		
1741	ACCAGCCAAATCATACACCGGACCTGAACAGGCCCAATATTTCTGCTCTATGAGGATGGC	1800
DB		
1801	CATGCTGTGGTGGCAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGAGTAGAACAAC	1860
QY		
1801	CATGCTGTGGTGGCAGATTTTGAGAAATCAAGATTTCTGAGTCCCTGGATGAAGACAAC	1860
DB		
1861	ATGACAAACCAACTCTGGGAACCTCCGTTGGATGGCTCTTGAGGTGTTTCACGCACTGCACT	1920
QY		
1861	ATGACAAAGCAGCAGGGAACCTCGCTGGATGGCCCTGAGGTGTTTCACACAGTGCACG	1920
DB		
1921	CGGTACACCAATCAAGCAGATGTCTTCAGCTATGTCTGTGTCTGTGGGAAATTTCTCACT	1980
QY		
1921	AGATACACCAATCAAGGCTGATGTCTTCAGTTACTCCCTGTGTCTGTGGGAGCTCCTCACT	1980
DB		
1981	GGCGAAATTCATTCGCTCATCTCAAGCAGCGGCTGGGCGAGCAGACATGGCTTACCAC	2040
QY		
1981	GGAGAAATTCATTCGCTCATCTCAAGCAGCGGCTGGCAGCAGAGATATGGCGTATCAC	2040
DB		
2041	CACATCAGACCTCCCATGGCTATTCATTTCCCAAGCCCATCATCTCTCTGATACGA	2100
QY		
2041	CACATCAGACCGCCCATCGGCTATTCATCCCAAGCCCATCTCTCTGATACGG	2100
DB		
2101	GGGTGGAAACGATGTCCTGAAGGAAGACCGGAATTTTCTTGAAGTTGTCTATGAAGTTAGAA	2160
QY		
2101	GGCTGGAATGATGTCTGAAGACACACAGAGTTCTCTTGAAGTGGTTAGCAAACTGGAG	2160
DB		
2161	GAGTGTCTTGAACATTTGAGCTCATGTCTCTGATCAAGTAAACAGCAGTGGGTCTCTC	2220
QY		
2161	GAGTGCCTTATCAATGTGGAGCTCATGTCTCCAGCATCAAGTAAACAGAGTGGCTCTCTG	2220
DB		
2221	TCACCTCTCTTCTTCTTGATTTGCTTGGTGAACCGGGGAGGACCTTGGCGGAGTCATGTG	2280
QY		
2221	TCACCTTCTCTTCTTCGATTCGCTGTGAGCCGGAGGCGCTTGGCGGAGCCACGTG	2280
DB		
2281	GCAGCATTAAGAGTGTGTTGGAATTTGGAATATGTCTTAAATGCAAGTCTCTATGCTGCT	2340
QY		
2281	GCAGCCTTACGAGCCGTTTTGAGTTGGAGTATGCGCTAAATGCAAGTCTCTATGCTGGG	2340
DB		
2341	TTGTCCCAAGTGTGGACATATTTCTCTCAAGTCTGTCTTTGGAGGAGATGAAAAGA	2400
QY		
2341	TGGTCCCAAGTGTGGAAACACATCTTAATTCGGGGCTGTCTTTGGAGGAGTGAATAGG	2400
DB		
2401	AGTCTTCAATACACACCCATTTGCAAAATATGGTATGTATCCGATCCCATGAGCTCAATG	2460
QY		
2401	AGCACCCAGTATTCAACTGTTGACAAATACGGCTATGTGTCTGATGCCATGAGCCTGACG	2460
DB		
2461	CATTTTCAATTTGCCGAAATPATGTAGCAGCTTTTGAGGACAGCAGC	2505
QY		
2461	CACCTTCACTTCCGCGCAGACGACAGCAACTTTTGAGGACAGCAAC	2505
DB		

RESULT 11				
AR269114	AR269114	Sequence 9 from patent US 6500654.	2505 bp	DNA
LOCUS	AR269114			
DEFINITION	Sequence 9 from patent US 6500654.			
ACCESSION	AR269114			
VERSION	AR269114.1	GI:29699957		

KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 2505)
AUTHORS	Raju, J.
TITLE	CARK protein and nucleic acid molecules and uses therefor
JOURNAL	Patent: US 650654-A 9 31-DEC-2002;
FEATURES	Location/Qualifiers
source	1..2505
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
Query Match	75.5%; Score 1892.2; DB 6; Length 2505;
Best Local Similarity	84.7%; Pred. No. 0;
Matches 2122; Conservative	0; Mismatches 383; Indels 0; Gaps 0;
Qy	1 ATGGGAATTATAATCTAGACCAACCCAACTTGTACTGATGAATGGAGAAAAGTC 60
Db	1 ATGGGAATTACAAATCCAGACCAACACAGACTTGTTCATGAATGGAGAAAAGTT 60
Qy	61 AGTGATCATATGTATACAAATAGAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA 120
Db	61 AGTGATCTTACGCTATTATCATAGAAAGGCTGGAGGATAACCTGCAGATCAAGGAAAA 120
Qy	121 GAACCTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180
Db	121 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCTTCAGTGAAGTCAATTTA 180
Qy	181 AATTACCGCATGAAATGGCTGTCTCTACTTCATTTATGTGCAATTTGTGGAGGCAAG 240
Db	181 AATTACCGCACAGAGCGTGGCTGTCCCTGCTACACCTCTGTGTCTGTGTGGGCAAC 240
Qy	241 AAATCACATATTCGAATCTTATGTTGAAAGGCTCGCCCATCTCGACTGACAGAAAT 300
Db	241 AAGTCACATATCGTGGCTTATGTTAAAGGCTCGTCCATCCAGATGACGAGAAAT 300
Qy	301 GGATTTACGCTTGCATTTAGCAGTTTACAGGATAATGACAGAAATTTGATCACTTCTGTG 360
Db	301 GGGTTTCCAGCTCTGCACCTGGCGCTTTACAGGACAGCCCGGAACCTTATCACTTCACTG 360
Qy	361 CTTACAGTGGAGTGATACAGAGGTTGATAGCGTGGCTCACTGCCCTCCATATT 420
Db	361 TTGCACAGGAGAGATGTTACAGAGTGGGATACGGTGGCTCAAGCCCTCCACATA 420
Qy	421 GCTCAATAGCTGGCCACCTAGAGGCTGTGATGTGCTGTGTGCAACATGGAGCTAATGTC 480
Db	421 GCTGCAATAGCTGGACACCCAGAGGCTGCAGAAAGTGTCTGTACAACATGGGCCCAATGTG 480
Qy	481 AATATTCAAGTACGAGTTTTTTTCACTCCATTTGCATTTGTCAGCGTACTATGGATGAA 540
Db	481 AATGTTCAAGATGCGGCTTTCTTCACTCCATTCGCAATTTGCAGCGTACTATGGCAGAG 540
Qy	541 CAGGTAACCTGCGCTTCTTTGAAATTTGGTGTGATGATGATAAATGTAAGTGGTGAAGTGA 600
Db	541 CAGGTAACCAAGTGTCTTTTGAAGTTTGGTGTGATGTCATGTAAGCGGTGAAGTGGG 600
Qy	601 GATGACCCCTCCACTAGCATCTGCAAAAGGATTTTGAATTTGCAAAATCTTTGATG 660
Db	601 GACAGGCTCTGCACCTGGCTCTGCAAAAGGCTTTCTTCAACATTTGTGAAATCTCTGTA 660
Qy	661 GAAGAGGAGCAAGCAGATGTAATGTCTCAAGATAATGAAGACCATGTCCCATCCAT 720
Db	661 GAAGAGGAGCAAGCAGATGTGAACGCTCAGGACAAATGAAGACCATGTCCCTCTGCAC 720
Qy	721 TTCTGTTCTCGATTTGGAACCAATGATATAGTATGATCTGCTGCAAGATGATTTGAA 780
Db	721 TTCTGTTCTCGATTTGGAACCAATGATATAGTATGATCTGCTGCAAGATGATTTAG 780
Qy	781 GTTCAACCTCATGTGTTAATATCTATGGAGATACCCCTTACCTTGCATGTGCTACAA 840
Db	781 GTCCAGCCTCACGTCTAATACATCTATGTTGACATCTCTTTGCACTTGTGATGCTTACANT 840
Qy	841 GGCAAAATTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
Db	841 GGAAATTTTGAAGTTGCCAAGGAAATTTGTCAGGTAACAGGAATCTGAAAGTCTGACTAAG 900
Qy	901 GAAACAATCTTTCAGTGAACAGCTTTTTCATAGTGTCTGTACCTATGGCAAGAGCATTTGAC 960
Db	901 GAAACAATCTTTCAGCGAGACAGCTTTTTCAGAGTCTGTACCTATGGCAAGAGCATTTGAC 960
Qy	961 CTAGTCAAAATTTCTTCTGTATGAGAATGTCAATAACATCAACACCAAGAAAGGATGGG 1020
Db	961 CTGTCAAAATTTCTTCTGTATCAGAATGCTGTGCAACATTTAACACCGAGGAAGATGGG 1020
Qy	1021 CACACTGGAATTACACTCTGCTTACACAGGTCACATTCGCTCGCTGGTTCAGTTCCTTACTG 1080
Db	1021 CACACAGATTGCACTCTGCTTCTACACGGCCATATCCGCTGGTTCAGTTCCTTACTT 1080
Qy	1081 GATAATGGAGCTGATATGAATCTTAGTGGCTTTGTGATCCAGCAGGTCTAGTGGTGA AAAA 1140
Db	1081 GATAATGGTGCAGATATGAATCTTGTGCTTGTGATCCAGCAGGTCTAGTGGTGA AAAA 1140
Qy	1141 GATGAGCAGACATGTTTGTATGTTGGCTTATGAAAAGGSCATGATGCCATTTGTACACTC 1200
Db	1141 GATGAGCAGACATGTTTGTATGTTGGCTTACGAGAAGGACATGATGCCATTTGTACACTC 1200
Qy	1201 CTGAAGCATTTAAGAGACACAAAGATGAATTCCTGTATGAATATTTCTCAGCTCGGA 1260
Db	1201 CTGAAGCCTTACAAGAGACCCAGAGGAGGTGCCATGTACGAATATTTCCAGCTCGGA 1260
Qy	1261 GGAGATGGCTCTATGTGTCTGTTCATCACCTTTGGGGAGAGATTAAAGCATGACAAAA 1320
Db	1261 GGAGATGGCTCTATGTGTCTGTTCCTTCCCCCTTTGGGCAAGATTAAAGCATGACAAAA 1320
Qy	1321 GAGAAGCAGATATTCCTCTCTAGAGCTGATTTGCCCTTTCATTTCCATTTCTCAGCTC 1380
Db	1321 GAGAAGCAGATGTTCTCTCTCTGAGGCTGAATCTACCTCCCGCTTCCATCTCCAATC 1380
Qy	1381 TCAGAAATGAGTTCCATGAGATTTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
Db	1381 TCCGAAATCGAGTTCCAGAGATTTACGGCTCGGCTTCTTTGGGAAAGTCTATAAGGG 1440
Qy	1441 CGATGAGAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTACTGTCTCCAAG 1500
Db	1441 CGATGAGAAATAAATAGTGGCTATAAAGCTTATCGAGCAATACCTACTGTCTCCAAG 1500
Qy	1501 TCAGATGTGGATATGTTTGGCGAGAGGTGTCATTTCTGCGAGCTCAATCATCCCTGTC 1560
Db	1501 TCAGACGTGGATATGTTTGGCGAGAGGTGTCATTTCTGCGAGCTCAACCCCTGTC 1560
Qy	1561 GTAAATTCAGTTTGTGGTGCTTGTCTGATGATCCAGCCAGTTTGCATTTGTCACTCAA 1620
Db	1561 GTGGTTTCAGTTTGTGGTGCTTGTCTGATGATCCAGCCAGTTTGCATTTGTCACTCAG 1620
Qy	1621 TACATATCAGGGGTTCTCTGTTCTCCCTCTCATGAGCAGAGAGGATTTCTTGAATTTG 1680
Db	1621 TACATTTACAGAGGCTCCCTGTTCTCCCTGCTTCTCATGAAGAGAGATTTCTTGAATTTG 1680
Qy	1681 CAGTCTAAATTAATTTATGTCAGTATGTTGTCGAAAGGATGAGTACCTTTCACAACTG 1740
Db	1681 CAGTCTAAATTAATCAATTTGCGGTAGACGTTGTCGAAAGGATGAGTACCTTTCACAGTTG 1740
Qy	1741 ACACAGCCAAATTAATCATCTGCTGACTTGAACAGTCAATATTTCTTCTATGAGGATGGG 1800
Db	1741 ACCAGCCAAATCATACACGCGACCTGAACAGCCACATATTTCTTCTATGAGATGGC 1800
Qy	1801 CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
Db	1801 CATGCTGTGGTGGCAGATTTTGGAGAAATCAAGATTTCTGCAAGTCCCTGGATGAAGACAAC 1860
Qy	1861 ATGACAAAACACCTGGGAACCTCGTTGGATGGCTCTGAGGTTTTCAGCAGTGCAC 1920
Db	1861 ATGACAAAGCACCCAGGAAACCTCGCTGGATGGCCCTGAGGTGTTTCAACAGTGCACG 1920

121	Qy	GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTTCAGTAAAGTCAATTTTA	180
121	Db	GAATTTCAAGAACTAAAGGCACATCTTTGGCTCTGATGAAGCCCTTCAGTGAAGTCAGTTTA	180
181	Qy	AAATTACCGCACTGAAAATGGGTGTCTCTACTTTCATTTATGTGTGATTTGTGGAGGCAAG	240
181	Db	AAATTACCGCACAGAGCGTGGCTGTCCCTGCTACACCTCTGCTGTGTCTGTGGGGCAAC	240
241	Qy	AAATTCACATATTCGAACCTCTTATGTTGAAAAGGGCTCCGCCCATCTCGACTGACAGAAT	300
241	Db	AAGTCACATATCCGTGCGCCCTTATGTTTAAAGGGCTCCGTCCATCCAGACTCAGCAGAAAT	300
301	Qy	GGATTTACAGCCTTTGCANTTTAGCAGTTTACAAAGGATAATGCAGAAATTGCATCACTTCTCTG	360
301	Db	GGGTTTCCAGCTCTGCACCTGGCCGCTTTACAGAGCACGCCGGACCTTATCACTTCACTG	360
361	Qy	CTTCACAGTGGAGCTGATATACACGAGGTTGATACGGTGGCTCACTGCCCTTCCATATT	420
361	Db	TTGCACAGCGAGCAGATGTTTCAGCAAGTGGGATACGGTGGCTTCACAGCCCTCCACATA	420
421	Qy	GCTCAATAGCTGGCACCCAGAGGCTGTGATGTGCTGTGCAACATGGAGCTAAATGTC	480
421	Db	GCTGCAATAGCTGCACACCCAGAGGCTGCAGAAAGTGTCTTACAACATGGGCCCAATGTG	480
481	Qy	AATATTCAAGATGCAGTTTTTTTCACTCCATTGCATATTGCAGGTACTATGGACATGAA	540
481	Db	AATGTTCAAGATGCGCTTCTTTCACCCCACTGCACATTTGCAGCCTACTATGGGCAOCAG	540
541	Qy	CAGGTAACCTGCCTCTTTTGAATTTGGTGTGATGTAAATGTAAGTGGTGAAGTTTGGGA	600
541	Db	CAGGTAACCAAGTGCTCTTTTGAAGTTTGGTGTGATGTCAATGTAAAGCGGTGAAGTTGGG	600
601	Qy	GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATATTGTCAAAAACCTCTTGATG	660
601	Db	GACAGGCTCTGCACCTGGCTCTGCAAGGGCTTCTTCAACATTTGTGAATCTCCTGGTA	660
661	Qy	GAAAGGCGAGCAAGCAGATGTGAATGTCTAAGATAATGAAGACCAATGTCCCCTCCAT	720
661	Db	GAAAGGCGAGCAAGCAGATGTGAACGCTCAGGACAATGAAGACCAAGTCCCTCTGCAC	720
721	Qy	TTCTGTTCTCGATTTGGACACCATGATAGTTAAGTATCTCTGCTGCAAAAGTGAATTTGAA	780
721	Db	TTCTGTTCTCGATTTGGACACCAATATAGTGAAGTACTCTGCTCCAGAGTGACTTAGAG	780
781	Qy	GTTCACACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACTGCGATGTCTACAT	840
781	Db	GTCCAGCCTCACGTCATTAACATCTATGGTGACACTTCTTGCACCTGCGATGTCTACAT	840
841	Qy	GGCAAAATTGAGTTGGCCAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG	900
841	Db	GGAAATTTGAGTTGGCAAGGAAATTTGTCCAGTTAACAGGAATCTGAAGTCTGACTAAG	900
901	Qy	GAAGAACATCTTTCAGTGAAAACAGCTTTTTCATAGTGTCTGTACCTATGGCAAGACATTCAC	960
901	Db	GAAGAACATCTTTCAGGAGACAGCTTTTTCACAGTGTCTGTACCTATGGCAAGAACATTCAC	960
961	Qy	CTAGTCAAAATTTCTTCTTGATCAGATGTCTAATAACATCAACCCAGGAGGATGGG	1020
961	Db	CTGTGTCAAAATTTCTTCTTGATCAGAAATGCTGTGAACATTTAACCCAGGAGGAGATGGG	1020
1021	Qy	CACACTGGATTTACACTCTGCTGTCTACACGGTTCACATTCGGCTGGTTCAAGTTCTTACTG	1080
1021	Db	CACACAGAAATTGCACTCTGCTGTCTACACGGCCATATCCGGCTGGTTCAAGTTCTTACTT	1080
1081	Qy	GATAATGGAGCTGATATGAATTCATAGTGGCTTTGATGCCAGCAGGTCTATGGGTGAAAAA	1140
1081	Db	GATAATGGTGAGATATGAATTCCTTGTGGCTTTGATGCCAGCAGGTCTATGGGTGAAAAA	1140
1141	Qy	GATGACAGACATGTTTGATGTGGGCTTATGAAAAAGGCGATGATGCCATTTGTCACTC	1200
1141	Db	GATGACAGACATGTTTGATGTGGGCTTACGAGAAAGACATGATGCCATTTGTTCATCTC	1200
1201	Qy	CTGAAAGCATTTAAGAGACCAACAGATGAAATTCGCCCTGTAATGAATATTCAGGCTCGA	1260

721	TTCTGTTCTCGAATTTTGACACCAACAATATAGTGAGCTACCTGCTCAGAGTGACTTAGAG	780
781	GTTCAACCTCACTGTTGTTTAATATCTATGAGATACCCCTTACACTGGCATGCTACAAAT	840
781	GTCCAGCGCTCAGTCATTAACATCTATGGTGACACTCTTTGCATCTGGCATGCTACAAAT	840
841	GGCAAAATTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG	900
841	GGAAATTTTGAAGTTGCCAAGGAAATTTGTCAGGTAACAGGAACCTGAAAGTCTGACTAAG	900
901	GAAACAATCTTTCACTGAAACAGCTTTTTCATAGTGCTTTGTAACCTATGGCAAGAGCAATGCAC	960
901	GAAACAATCTTTCACTGAAACAGCTTTTTCATAGTGCTTTGTAACCTATGGCAAGAGCAATGCAC	960
961	CTAGTCAAAATTTCTTCTTTGATCAGAATGTCATAAACATCAACACCAAGGAAGGATGGG	1020
961	CTGGTCAAAATTTCTTCTTTGATCAGAATGTCATAAACATCAACACCAAGGAAGGATGGG	1020
1021	CACACTGGAATTTACACTCTGCTTGGCTACCAAGGTCACATTCGGCTGGTTTCAGTTCTTACTGT	1080
1021	CACACAGAAATTCGCACTCTGCTTGGCTACCAAGGTCACATTCGGCTGGTTTCAGTTCTTACTGT	1080
1081	GATAATGGAGCTGATATGAATCTAGTGGCTTTGTGATCCACGACAGGTCCTAGTGGTGAAAAA	1140
1081	GATAATGGTGACAGATATGAATCTTGTGGCTTTGTGATCCACGACAGGTCCTAGTGGTGAAAAA	1140
1141	GATGACGACAGATGTTTGAATGCTGGGCTTATGAAAAGGCGCATGATGCCAATTCGCACATC	1200
1141	GATGACGACAGATGTTTGAATGCTGGGCTTATGAAAAGGCGCATGATGCCAATTCGCACATC	1200
1201	CTGAAGCAATTAAGAGACCAACAAGATGAAATTCGCCCTGTAATGAATATTCTCAGCGCTGGA	1260

Db 104 AGTGAATCTTACGCTATATATATAGAAAGGCTGGAGGATACCTGCAGATCAAGAAAT 163
Qy 121 GAAGTGAAGAACTAAGAAATATATTTGGCTCTGATGAAGCCTTCAGTAAGTCAATTTA 180
Db 164 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCTTCAGTGAAGTCAAGTTA 223
Qy 181 AATTACCGCACTGAAAATGGGCTGCTCTACTTCAATTTATGTTGATTTGGAGGCAAG 240
Db 224 AATTACCGCAAGAGGGTGGCTGCTCTGCTGTACACCTCTGCTGTGCTGGGGCAAC 283
Qy 241 AAATCAATATTTCSAACTTTATATGTTGAAGGGCTCCGCCCATCTCGACTGACAAAGAAAT 300
Db 284 AAGTCAATATCCGTGCTCCCTTATGTTAAAGGGCTCCGTCCATCCAGACTGACGAGAAAT 343
Qy 301 GGNATTACAGCCTTGCAATTTACAGTTTACAGGATATGAGAAATGATGATCACTTCTCTG 360
Db 344 GGGTTTCCAGCTCTGCACTGCGCTTTTACAGGACAGCCCGGAACCTTATCACTTCACTG 403
Qy 361 CTTCAAGTGGAGCTGATATACAGCAGTTGATACGGTGGCTCACTGCCCTCCATATT 420
Db 404 TTGCAAGCGGAGCAGATGTTTACAGAAAGTGGATACGGTGGCTTACAGCCCTCCACATA 463
Qy 421 GCTAAGTGGCGCACTAGAGGCTGCTGATGCTGTTGCAACATGAGGCTAAATGTC 480
Db 464 GCTGCAATAGCTGGACACCCAGAGGCTGAGAAAGTGTCTGCTACAAATGGGGCCAAATGTG 523
Qy 481 AATATTTCAAGATGCACTTTTTCCTCATTTGCTATATGAGCGCTACTATGCAATGAA 540
Db 524 AATGTTCAAGATGCGCTTCTTCCACCACTGACATGAGCTACTATGGGCACGAG 583
Qy 541 CAGGTAACCTCGCTTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGGA 600
Db 584 CAGGTAACCACTGCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTTGGA 643
Qy 601 GATAGACCCCTCCACCTAGCATCTGCAAGGATTTTGAATATGCAAACTCTTGTATG 660
Db 644 GACAGGCTCTGCACTGCGCTCTGCAAGGCTTCTTCAACATTTGGAATCTCTGCTGTA 703
Qy 661 GAAGAAAGGAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720
Db 704 GAAGAAAGGAGCAAGCAGATGTAAGCTCAGGACCAATGAAGACCATGTCCCTCTGCAC 763
Qy 721 TTCTGTTCTGATTTGGACCACTGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 764 TTCTGTTCTGATTTGGACCACTGATGATGATGATGATGATGATGATGATGATGATGAT 823
Qy 781 GTTCAACCTCATGTTGTAATATCTATGAGATACCCCTTACACCTGGCATGCTACAAAT 840
Db 824 GTCCAGCTCAGTCAATTAACATCTATGATGATGATGATGATGATGATGATGATGATGAT 883
Qy 841 GGCAAAATTTGAAGTTGCCAAGGAATCATCCAAATATACAGGAACAGAAAGTCTGACTAAG 900
Db 884 GGAAATTTTGAAGTTGCCAAGGAATTTGTCAGGTAAACAGGAATCTGACTAAG 943
Qy 901 GAAACATCTTCAGTGAACAGCTTTTCACTGATGATGATGATGATGATGATGATGATGATGAT 960
Db 944 GAAACATCTTCAGGAGAGAGCTTTTCACTGATGATGATGATGATGATGATGATGATGATGAT 1003
Qy 961 CTAGTCAAAATTTCTTCTGATCAGATGCTATAAAACATCAACCAACAGGAAGGATGGG 1020
Db 1004 CTGGTCAAAATTTCTTCTGATCAGATGCTGTGAACATTAACCAACAGGAAGGATGGG 1063
Qy 1021 CACATGGATATCACTGCTGCTTACAGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1064 CACAGGATTTGACCTCTGCTTCTTCAACAGGCTATCCGCTGCTGCTGCTGCTGCTGCTGCTG 1123
Qy 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCCAAGCAGGCTTCTAGTGGTGAATA 1140
Db 1124 GATAATGGTGCAGATATGAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1183
Qy 1141 GATGAGCAGACATGTTTGTGATGCTGGGCTTATGAAAGGGGCTGATGCCATTTGTCACATC 1200

Db 1184 GATGAGCAGACATGTTTGTGATGGGCTTTACAGAAAGGACATGATGCCATTTGTACACTC 1243
Qy 1201 CTGAAGCATTTAAGAGACCAAGATGAATTTGCCCTGTATGAATATTTCTCAGCCCTGGA 1260
Db 1244 CTGAAGCATTTAAGAGACCAAGATGAATTTGCCCTGTATGAATATTTCTCAGCCCTGGA 1303
Qy 1261 GGAGATGGCTCTCTATGCTGCTGTTTCCATCAACCTTTGGGGAAGATTTAAAGCATGACAAA 1320
Db 1304 GGAGATGGCTCTCTATGCTGCTGTTTCCCTTCCCTTGGGCAAGATTTAAAGCATGACAAA 1363
Qy 1321 GAGAGGAGGATATTTCTCTCTAGAGCTGATGCTTCCATTTCCATTTCTTCCATTTCTCAGCTC 1380
Db 1364 GAGAGGAGGATGTTTCTCTCTGAGGCTGAATCACTCCCTCCCTTCCATTTCTCAGCTC 1423
Qy 1381 TCAGAAATTTGAGTTTCCATGAGATTTATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
Db 1424 TCCGAAATTCGAGTTCCACGAGATTTATCGGCTCGGTTCTTTTGGGAAAGTATATAAGGG 1483
Qy 1441 CGATCGAGAAATATAATAGTGGCTATAAAAGTTTATCGAGCCATACCTACTGCTCCAAAG 1500
Db 1484 CGATCGAGAAATATAATAGTGGCTATAAAAGTTTATCGAGCCATACCTACTGCTCCAAAG 1543
Qy 1501 TCAGATGAGGATATGTTTGGCGAGGCTGCTTCTCTGCCAGCTCAATCATCCCTGTC 1560
Db 1544 TCAGAGCTGGATATGTTTGGCGAGGCTGCTTCTCTGCCAGCTCAACCCCTGTC 1603
Qy 1561 GTAAATTCAGTTTGGGCTGCTTGAATGATCCAGCAGTTTGGCATTTGCTTACCTCAA 1620
Db 1604 GTGGTTCAGTTTGGGCTGCTGCTGCTGATGACCCAGTCACTTTGCTTGTCTCAGT 1663
Qy 1621 TACATATCAGGGGTTCTCTGCTTCTCTCTCTCATGAGCAGAGAGATTTCTTGAATTTG 1680
Db 1664 TACATTTAGGAGGCTTCTCTGCTTCTCTCTCTCATGAGCAGAGAGATTTCTTGAATTTG 1723
Qy 1681 CAGTCTAAATTAATTTAGCAGTATGCTTGCCTAAAGGCTATGAGTACCTTTCAACACTG 1740
Db 1724 CAGTCTAAATTAATCATTTGGGTAGACCTTGCCTAAAGGCTATGAGTACCTTTCAACACTG 1783
Qy 1741 ACACAGCCTAATTAACATGCTGCTGCTTGAAGTCACTAATTTCTTCTATGAGAGTGG 1800
Db 1784 ACCAGCCTAATTAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1843
Qy 1801 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1844 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1903
Qy 1861 ATGACAAAACACTGGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1904 ATGACAAAACACTGGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1963
Qy 1921 CGGTACACCATCAAGCAGATGCTTTCAGCTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 1964 AGATACCATCAAGCAGATGCTTTCAGCTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2023
Qy 1981 GGCGAAATTTCCATTTCCCTCATCTCAAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db 2024 GGAGAAATTTCCATTTCCCTCATCTCAAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2083
Qy 2041 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCATATCATCTCTGCTGATAGA 2100
Db 2084 CACATCAGACCTCCCATTTGGCTATTTCCATTTCCCAAGCCCATCTCATCTCTGCTGATAGG 2143
Qy 2101 GGGTGAAAGCAGATGCTTCTGAAAGGAAGCCGAAATTTCTGAAGTTTGTCTGAAGTTAGAA 2160
Db 2144 GGCTGAAATGCTATGCTTCTGAAAGGAGCAGCAGAGTTTCTCTGAAGTCTGTAGCAAACTGGAG 2203
Qy 2161 GAGTGTCTCTGCAACATTTAGAGCTGATGCTCTCTGCTCAAGTAAAGCAGAGTGGCTCTCTC 2220
Db 2204 GAGTGTCTCTGCAACATTTAGAGCTGATGCTCTCTGCTCAAGTAAAGCAGAGTGGCTCTCTC 2263
Qy 2221 TCACCTTCTTCTTCTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db 2264 TCACCTTCTTCTTCTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2323

QY	2281	GCAGCATTAAGAAGTCGTTTGGAAATGGAATATGCTCTAAATGCAAGGTCCTATGCTGCT	2340	Db	181	GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCCTTCAGTGAAGTCAGTTTA	240
Db	2324	GCAGCCTTACGGAGCGTTTGGAGTGGAGTATGCCCTAAATGCAAGGTCCTATGCTGGG	2383	QY	181	AATTTACCGCACTGAATGGCTGCTCTACTTCTATTTATGTTGATTTGTCGAGGCAAG	240
QY	2341	TGTCCCAAGTGTGGAGCAATATCTCTCAAGGTCTGTCTTTGGAGGAGATGAAAAA	2400	Db	241	AATTTACCGCAAGAGCGTGGCTCTCCCTGTCTACACCTCTGCTGTGTCTGTGGCGCAAC	300
Db	2384	TGGTCCCAAGTGTGGAGCAACACTCTAATCCGGGCTGTCTTTGGAGGAGATGAAATAGG	2443	QY	241	AAATACATATTCGAACTCTTATGTTGAAAGGGCTCCGCCCATCTCGACTGACAGAAAT	300
QY	2401	AGTCTTCAATACACACCCATTGACAAATATGGCTATGTATCCGATCCCATGAGCTCAATG	2460	Db	301	AAGTACATATCCGTGCCCTTATGTTAAAGGGCTCCGTCCATCCAGACTGACGAGAAAT	360
Db	2444	AGCACCCAGTATTCACTGTGTGACAAATACGGCTATGTCTGATCCCATGAGCCCTGACG	2503	QY	301	GGATTTACAGCCTTGCATTTAGCAGTTTACAGGTAATGACAGATTCATCTCTCTG	360
QY	2461	CATTTTCATTTCCCGAAATAGTAGCAGCTTTGAGGACAGCAGC	2505	Db	361	GGGTTCAGCTCTGCACCTGGCCGTTTACAAGGACAGCCCGAACTTATCATCTTCACTG	420
Db	2504	CACCTTACTCCGCCAAGACGACGACCAACTTTGAGGACAGCAAC	2548	QY	361	CTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCCCTCACTGCCCTCCATATT	420
RESULT 14				Db	421	TTGCACAGCGGACAGATGTTCAGCAAGTGGATACGGTGGCTTACAGCCCTCCACATA	480
LOCUS	BD269561			QY	421	GCTACAAATAGCTGGCCACCTAGAGGCTGCTGATGTGCTTGTGCAACATGGAGCTTAATGC	480
DEFINITION	Novel CARP protein and nucleic acid molecules and uses therefor.			Db	481	GCTGCAATAGCTGGACACCCAGAGGCTGAGAAGTGTCTGCTACAACTATGGGCCCAATGTG	540
ACCESSION	BD269561			QY	481	AAATATTCAAGATGCAAGTCTTTTCACTCCATTCATATTGAGCGCTACTATGGACATGAA	540
VERSION	BD269561.1	GI:33079329		Db	541	AATGTTCAAGATGCCGTCTTCTTCACCCACCTGACANTTGCAGCTTACTATGGGACGAG	600
KEYWORDS	JP 2002535962-A/3.	(Norway rat)		QY	541	CAGGTAACTCGCCTCTTTTGAATTTGGTGTCTGATGTAATGTAAGTGGTGAAGTTGGA	600
SOURCE	Rattus norvegicus (Norway rat)			Db	601	CAGTAACAGTGTCTCTTTTGAAGTTGGTGTGATGTCAATGTAAGCGGTGAAGTTGGG	660
ORGANISM	Rattus norvegicus			QY	601	GATAGACCCCTCCACTAGCATCTGCAAAAGGATTTGTAATATGCAAACTCTTCTGATG	660
REFERENCE	1	(bases 1 to 3026)		Db	661	GACAGGCTCTGACCTGGCCCTCTGCAAGGGCTCTTCAACATTTGTAAGTCTTCTGCTA	720
AUTHORS	Raju, J.			QY	661	GAAGAAGGAGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT	720
TITLE	Novel CARP protein and nucleic acid molecules and uses therefor			Db	721	GAAGAAGGAGCAAGAGCAGATGTGAACGCTCAGGACAATGAAGACCATGTCCCTCTGCAC	780
JOURNAL	Patent: JP 2002535962-A 3 29-OCT-2002;			QY	721	TTCTGTCTCGAATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAGTGAATTTGGA	780
COMMENT	MILLENNIUM PHARMACEUTICALS INC			Db	781	TTCTGTCTCGAATTTGGACACCAATATAGTTAAGTATCTGCTGCAAGTGAATTTGGA	840
	OS Rattus norvegicus (rat)			QY	781	GTTCACACCTCATGTTGTTAAATATCTATGAGATACCCCTTTACACCTGGCATGTACAAT	840
	PN JP 2002535962-A/3			Db	841	GTCCAGCTCAGCTCATTAACATCTATGTTGACACTCTTTGACCTGGCATGTACAT	900
	PD 29-OCT-2002			QY	841	GGCAATTTGAAATTTGCCAAGGAAATTTGTCAGGTAATCCAGGAACCTGAAAGTCTGACTAAG	960
	PF 10-DEC-1999	JP 2000586772		Db	901	GGAAATTTGAAATTTGCCAAGGAAATTTGTCAGGTAATCCAGGAACCTGAAAGTCTGACTAAG	960
	PR 11-DEC-1998	US 60/111938, 14-APR-1999	US 09/291839	QY	901	GAAACATCTTCAGTGAACACAGCTTTTTCATAGTGTCTGTACTATGCAAGAGCATTTGAC	960
	PT JAYASEELAN RAJU			Db	961	GAAACATCTTCAGGAGACAGCTTTTTCATAGTGTCTGTACTATGCAAGAGCATTTGAC	1020
	PC C12N15/09, C07K16/40, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/10, C12Q1/02,			QY	961	CTAGTCAAAATTTCTTTGATCAGAAATGTCTAAACATCAACCAAGGAGGATGGG	1020
	PC C12Q1/68, G01N33/15, G01N33/53, G01N33/68//A61K38/45,			Db	1021	CTGGTCAAAATTTCTTTGATCAGAAATGCTGTGAACATTAACCAACCGAGGAGATGGG	1080
	PC A61K39/395,			QY	1021	CACACTGGATTAACACTGCTGTGTACACCGTCCATTTGCTGCTGCTTCTGCTTCTACTG	1080
	PC A61K45/00, A61P9/04, A61P35/00, A61P43/00, C12N15/00, PC C12N5/00,			Db	1081	CACAGGATTCGATCTGCTTTGTCTTACACCGCATATFCCGCTGGTTCAGTTCTCTACTT	1140
	PC A61K37/52			QY	1081	GATAATGAGCTGATATGAATCTAGTGGCTTTGTGATCCAGCAGGCTCTAGTGGTGAATAA	1140
	CC Novel CARP protein and nucleic acid molecules and uses			Db	1141	GATAATGAGCTGATATGAATCTAGTGGCTTTGTGATCCAGCAGGCTCTAGTGGTGAATAA	1200
	therefor			QY	1141	GATAATGAGCTGATATGAATCTAGTGGCTTTGTGATCCAGCAGGCTCTAGTGGTGAATAA	1200
	FH Key	Location/Qualifiers		Db	1141	GATAATGAGCTGATATGAATCTAGTGGCTTTGTGATCCAGCAGGCTCTAGTGGTGAATAA	1260
	FT CDS	(61), (2565).		QY	1141	GATAATGAGCTGATATGAATCTAGTGGCTTTGTGATCCAGCAGGCTCTAGTGGTGAATAA	1260
	source	1. 3026		Db	1201	CTGAAAGCTATTAAGAGACCAACAGATGAATTTGCCCTGTATGAATGAATTTCTCAGCCTGGA	1260
		/organism="Rattus norvegicus"		QY	1201	CTGAAAGCTATTAAGAGACCAACAGATGAATTTGCCCTGTATGAATGAATTTCTCAGCCTGGA	1320
		/mol_type="genomic DNA"		Db	1201	CTGAAAGCTATTAAGAGACCAACAGATGAATTTGCCCTGTATGAATGAATTTCTCAGCCTGGA	1320
		/db_xref="taxon:10116"		QY	1201	CTGAAAGCTATTAAGAGACCAACAGATGAATTTGCCCTGTATGAATGAATTTCTCAGCCTGGA	1320
				Db	1201	CTGAAAGCTATTAAGAGACCAACAGATGAATTTGCCCTGTATGAATGAATTTCTCAGCCTGGA	1320
FEATURES				QY	1201	CTGAAAGCTATTAAGAGACCAACAGATGAATTTGCCCTGTATGAATGAATTTCTCAGCCTGGA	1320
ORIGIN				Db	1201	CTGAAAGCTATTAAGAGACCAACAGATGAATTTGCCCTGTATGAATGAATTTCTCAGCCTGGA	1320
Query Match		75.5%	Score 1892.2;	DB 6;	Length 3026;		
Best Local Similarity		84.7%	Pred. NO. 0;				
Matches 212;	Conservative	0;	Mismatches 383;	Indels	0;	Gaps	0;
QY	1	ATGGGAATTTATAATCTAGACCAACCCAACTTGTACTGTGATGAATGGAAGAAAAAGTC	60				
Db	61	ATGGGAATTTATAATCTAGACCAACCCAACTTGTACTGTGATGAATGGAAGAAAAAGTC	120				
QY	61	AGTGAATCATATTGTTATCAATAGAAAGATTAGAAGATGACCTGCAGATCAAGAAAAA	120				
Db	121	AGTGAATCTTACGCTATTATCATAGAAAGGCTGGAGGATTAACCTGCAGATCAAGAAAAA	180				
QY	121	GAACTGACAGAACTAAGGAATATATTTGGCTCTGTATGAAGCCCTTCAGTAAAGTCAATTTA	180				

Qy	541	CAGGTAAC	CTGCCTCTCTTTT	GAAATTTTGGT	CTGATGTA	ATAATGTA	AGTGAAGTGG	TGAAGTTGG	600			
Db	601	CAGGTAA	CCAGTGTCTCTTTT	GAAAGTTTGGT	CTGATGTA	ATAATGTA	AGTGAAGTGG	TGAAGTTGG	660			
Qy	601	GATAGAC	CCCTCCAC	CTAGCATCTG	CAAAAAGG	ATCTTGA	ATAATTTG	CAAAAAC	CTCTTGATG	660		
Db	661	GACAGG	CTCTGCAC	CTGCCTCTG	CAAAAGG	CTTCTTCA	ATTTGTG	GAAC	CTCTGGTA	720		
Qy	661	GAAGAAG	CAGCAAAAG	CAGATGTGA	ATGCTCA	AGATAA	TAAAGACA	CAATGTC	CCCACTCAAT	720		
Db	721	GAAGAAG	GAGCAAAAG	CAGATGTGA	ATGCTCA	AGATAA	TAAAGACA	CAATGTC	CCCACTCAAT	780		
Qy	721	TTCTGTT	CTCGATTTTGG	ACACCATG	ATATAGTT	TAAGTATCTG	CTGCAAAAG	TGATTTG	GAA	780		
Db	781	TTCTGTT	CTCGATTTTGG	ACACCA	CAATATAGT	GAGTAC	CTGTCTG	CAAGTGA	CTTAGAG	840		
Qy	781	GTTCAC	CTCATGTTTGT	TTAATATCTAT	GTGAG	ATACCCCTTAC	CTGCGCATG	CTACAAT		840		
Db	841	GTCCAG	CTCACGTCA	TTTAA	CATCTAT	GTGTGAC	ACTCTTTG	CACTG	GCATGTCTACAAT	900		
Qy	841	GGCAAA	TTTGAAGTTTCC	AAAGAAATCAT	CTCA	AAATATC	AGGAAC	AGAAAGTCT	GACTAAG	900		
Db	901	GGAAA	TTTGAAGTTTCC	AAAGAAATTTG	CTCAG	GTAA	CAAGAACTG	AAAGTCT	GACTAAG	960		
Qy	901	GA AAA	CATCTTCAGT	GTAAACAG	CTTTTTCAT	GTGCTTTG	TACCTAT	TGCAAG	AGCATTTGAC	960		
Db	961	GA AA	CACTTC	CAGCAGAC	ACTTTTTCA	CAGTGTCT	TGTAC	CTATG	GCAGAA	CACTTGAC	1020	
Qy	961	CTAGTCA	AAATTTCTTCT	TGTATC	AGTCA	GTGATAA	CAATCA	CAACCA	AGAAAGG	ATGGG	1020	
Db	1021	CTGGTCA	AAATTTCTTCT	TGTATC	AGATG	CTGTG	AACTTAA	CAACCA	AGAAAGG	ATGGG	1080	
Qy	1021	CACATG	GAATAC	ACTCTG	TTGCTAC	ACAGGT	CACATTCG	CGTGTCT	CTTACTG		1080	
Db	1081	CACACAG	ATTG	CACTCTG	TTGCTAC	ACAGG	CCATTCGG	CTGTG	CTTACTT		1140	
Qy	1081	GATAAT	GAGAGT	GTATGA	ATCTAGT	GGCTTGTG	ATCCCAG	CAGGCTCTAGT	GGTGA	AAA	1140	
Db	1141	GATAAT	GGTGC	AGATATGA	ATCTTGT	CGCTTGTG	ATCCCAG	CAGGCTCTAGT	GGTGA	AAA	1200	
Qy	1141	GATG	CAGAC	ATGTTTGA	CTGGGCTTAT	GAAAGG	GCATGATG	CCATTTG	TCACACTC		1200	
Db	1201	GATGAG	CAGACATGTT	GATGTGG	CTTAC	GAAAGG	ACATGATG	CCATTTG	TACACTC		1260	
Qy	1201	CTGAAG	CAATTA	AGAGACCA	AGATGA	ATTCCT	GTAAATG	AAATTTCT	CAGCCTGGA		1260	
Db	1261	CTGAAG	CACTA	AGAGACCC	AGAGGAG	CTGCCAT	GTAA	CGAATTTCC	CAGCCTGGA		1320	
Qy	1261	GGAGAT	GGCTCCTAT	GTGCTGTTC	CACTAC	CCCTTGGG	GAAGATTA	AAAGCAT	CACAAA		1320	
Db	1321	GGAGAT	GGCTCCTAT	GTGCTGTTC	CACTCCT	CGGCTTGGG	CAAGATTA	AAAGCAT	CACAAA		1380	
Qy	1321	GAGAAG	GCAGATATTC	CTCCT	CTAAG	AGCTGGA	ATGCC	TTACATTTCC	ATCTT	CAGCTC	1380	
Db	1381	GAGAAG	GCAGATGTTCT	CCTCCT	CGGCTG	GA	ACTAC	CTCCGCTTCC	ATCTT	CCA	CTC	1440
Qy	1381	TCAGAAA	TTGATGTC	CAATGATTTGG	CTCA	GGTTCTTTTGG	AAAGTATATA	AAAGGA			1440	
Db	1441	TCGAAA	TCGAGTTCC	AGATTTG	GTGCTCA	GGTTCTTTTGG	AAAGTATATA	AAAGGA			1500	
Qy	1441	CGAT	GCAGAA	TATAA	ATAGTGG	CGATCA	ACGAG	CCAA	CACCTACT	GTCTCC	AG	1500
Db	1501	TCAGAT	GTGGATGTTTT	TGCG	AGAGGTTCC	ATTTCTCTG	CCAGCTCA	ATTCAT	CCCTGC		1560	
Qy	1561	TCAGAG	CTGGATATGTTTT	TGCG	AGAGGTTCC	ATTTCTCTG	CCAGCTCA	ATTCAT	CCCTGC		1620	
Db	1621	GTGGT	TCAGTTTGGG	TGCTGCT	GAATGATCC	AGGTTTGG	CAATTTG	CCATTTG	CACTCA		1620	
Qy	1621	TACATAT	CAGGGGTTCTCTG	TTCTC	CTCTCAT	GAGCAG	AGAGATTTCT	GATTTG			1680	

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-May-1999
C;Accession: T04683
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Jesse, T.; Heijnen, L.; Vos, P.; Mewes,
submitted to the Protein Sequence Database, July 1998
A;Reference number: Z15381
A;Accession: T04683
A;Molecule type: DNA
A;Residues: 1-553 <BEV>
A;Cross-references: EMBL:AL031135
A;Experimental source: Cultivar Columbia; BAC clone F8D20
C;Genetics:
A;Map position: 4
A;Introns: 69/2; 107/3; 176/2; 194/3; 231/3; 289/2; 325/2; 350/1; 408/3; 440/3; 467/3; 4
A;Note: F8D20.290

Query Match 10.8%; Score 473.5; DB 2; Length 553;
Best Local Similarity 29.1%; Pred. No. 4.2e-18;
Matches 134; Conservative 79; Mismatches 160; Indels 87; Gaps 15;

QY 353 IRUVQF---LLDNGADMLVADPPSSSEKDEQT-----CLMWAY 390
DB 103 VRCVQVSTLNGSGDV-----DPSDPAYNDAQSSYNSRLAPPTFGSSNFALTOAY 157
QY 391 EKGH---DAIVTLKHKYKRPQDELPCNEYSP-----GGD 422
DB 158 -KDAQDDSAVNAQLPNSRPWHEITFTIDRPKLLSQLTSMGLGNIQEAHAFSTAD 216
QY 423 GSXVSV-----PSPGKTKSMT-----KEKADILLRAGLPS-----HFHLQLSIE 464
DB 217 GFSLDVFFVVDGWSQEDQPCKSQKSIITFFBHDKSTNELLFACVEIPTDGTDEWIDMKQLK 276
QY 465 FHLIIGSGFGKYKGRCKNIKVAIKRYRANTYCSKSDVDM---PCREVSLCOLNHPVCV 521
DB 277 IEKKVACSGELFGFYCSQVAKILKP-----ERNVAMELRFSQVYINRKYRHKNV 332
QY 522 IQFVGACLNDFPSQFAIVTOYISGGSLFSLHHEQKRILDLQSKLIIADVAKGMEYLNLT 581
DB 333 VQFAGACTRSPN-LCIVTEPTMGSIYDFLHKHGVFKIQSLKVALDVSKGMNYLHQ-- 389
QY 582 QPIIHRDLNSHLLVDGHAVVADGESRFLOSLEDNMTKQGNLRWMADEVFTQCTR 641
DB 390 NNIHRDLKTNILMDSEHVVVADFGVARV--QTSSGVMTAETGYRWMADEV--EHKP 446
QY 642 YTIKADYFVALCLWEITLGEIPFAHLKPAAPAAADMAVHHIRPPIGYSIPKPISSLLIRG 701
DB 447 YDHRADVFAIVLWELLTGEIPSYLTPLOAVGVVQGLRKPKEPKETHPKLTLEK 506
QY 702 WNACPEGRPFSEVVMKLECLNIELMSPASSNSGSL 741
DB 507 WQDPAALRPNAETIEMNLQIREVIDLSLHKDKHGYS 546

RESULT 5
T18276
protein-tyrosine kinase (EC 2.7.1.112) 1 - slime mold (Dictyostelium discoideum)
A;Cross-references: Dictyostelium discoideum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000
C;Accession: T18276; A35670
R;Nuckolls, G.H.; Osherov, N.; Loomis, W.F.; Spudich, J.A.
Development 122, 3293-3305, 1996
A;Title: The Dictyostelium dual-specificity kinase splA is essential for spore different
A;Reference number: Z18852; MUID:97053827; PMID:8898241
A;Accession: T18276
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1584 <NUC>
A;Cross-references: EMBL:U32174; NID:g974333; PID:g974334; PIDN:ARB41125.1
R;Tan, J.L.; Spudich, J.A.
Mol. Cell. Biol. 10, 3578-3583, 1990
A;Title: Developmentally regulated protein-tyrosine kinase genes in Dictyostelium discoi
A;Reference number: A35670; MUID:90287147; PMID:1972546
A;Accession: A35670
A;Status: preliminary

A;Molecule type: mRNA
A;Residues: 'R',1249-1434, 'L',1436-1584 <TAN>
A;Cross-references: GB:M33785; NID:g167775; PIDN:AAA33202.1; PID:g167776
A;Note: the authors translated the codon TAT for residue 271519 as Thr
C;Genetics:
A;Gene: splA
A;Introns: 47/3; 72/2
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolc
C;Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F;1287-1566/Domain: protein kinase homology <KIN>
F;1295-1303/Region: protein kinase ATP-binding motif

Query Match 10.6%; Score 465; DB 2; Length 1584;
Best Local Similarity 37.7%; Pred. No. 4.2e-17;
Matches 109; Conservative 52; Mismatches 100; Indels 28; Gaps 6;

QY 454 SMFHLQLSIEFHEIIGSGFGKYKGRCKNIKVAIKRYRANTYCSKSDVDMFCEVSL 513
DB 1280 SEVEIDFNELEFGQTIGKGFGEVGRYWRDVAIKIYRDQFKTSKSLVNFQNEVIL 1339
QY 514 COLNHPCVIQFVGAC-LNDPSQFAIVTOYISGGSL-----FSLHHEQKRILDLQSKLI 565
DB 1340 SKURHENVVQFLOACTAGGEDHHCIVTEWGGSLRQFLTDHFNLEQNPHI-----RLK 1394
QY 566 IADVAKGMEYLNLTQPIIHRDLNSHLL-----YEDGHAVVADGESRF 612
DB 1395 LALDIAGKMNVLHGWTPTPIIHRDLSSRNILLDHNIDPKNPVVSRRQDIKCKISDFGLSR- 1453
QY 613 LQSLDEDNMTKQGNLRWMADEVFTQCTRITKADYFVALCLWEITLGEIPFAHLKPA 672
DB 1454 LKKEQASQMTQSGVCIYPNAPEVFKGDSN-SEKSDYISGVLFELLTSDRFPQDMKPMK 1512
QY 673 AAADMAVHHIRPPIGYSIPKPISSLLIRGNACPEGRPFSEVVMKLE 721
DB 1513 MAHLAAVESVRPIPLTTSSKNKEILTQCWDSPDSRPTFKQIIVHLKE 1561

RESULT 6
B35049
ankyrin 1, erythrocyte splice form 3 - human
N;Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N;Contains: ankyrin 2.2, erythrocyte
C;Species: Homo sapiens (man)
C;Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C;Accession: B35049
R;Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, D.; Cheung, M.C.; K
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A;Title: cDNA sequence for human erythrocyte ankyrin.
A;Reference number: A35049; MUID:90175370; PMID:1689849
A;Accession: B35049
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1856 <LAM>
C;Genetics:
A;Gene: GDB:ANK1; ANK
A;Cross-references: GDB:118737; OMIM:182900
A;Map position: 9p11.2-9p11.2
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>
F;2-1513,1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F;44-76/Domain: ankyrin repeat homology <AN01>
F;77-105/Domain: ankyrin repeat homology <AN02>
F;110-142/Domain: ankyrin repeat homology <AN03>
F;143-171/Domain: ankyrin repeat homology <AN04>
F;172-204/Domain: ankyrin repeat homology <AN05>
F;205-237/Domain: ankyrin repeat homology <AN06>
F;238-270/Domain: ankyrin repeat homology <AN07>
F;271-303/Domain: ankyrin repeat homology <AN08>
F;304-336/Domain: ankyrin repeat homology <AN09>
F;337-369/Domain: ankyrin repeat homology <AN10>
F;370-402/Domain: ankyrin repeat homology <AN11>
F;403-435/Domain: ankyrin repeat homology <AN12>

A:Molecule type: mRNA
A:Residues: 1-1862<RES>
A:Cross-references: GB:M84756; NID:g191939; PIDN:AAA37236.1; PID:g191940
C:Species: Dictyostelium discoideum
C:Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 04-Feb-2000
C:Accession: B35670
A:Gene: Ank-1
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:40-72/Domain: ankyrin repeat homology <AN01>
F:73-105/Domain: ankyrin repeat homology <AN02>
F:106-138/Domain: ankyrin repeat homology <AN03>
F:139-167/Domain: ankyrin repeat homology <AN04>
F:168-200/Domain: ankyrin repeat homology <AN05>
F:201-233/Domain: ankyrin repeat homology <AN06>
F:234-266/Domain: ankyrin repeat homology <AN07>
F:267-299/Domain: ankyrin repeat homology <AN08>
F:300-332/Domain: ankyrin repeat homology <AN09>
F:333-365/Domain: ankyrin repeat homology <AN10>
F:366-398/Domain: ankyrin repeat homology <AN11>
F:399-431/Domain: ankyrin repeat homology <AN12>
F:432-464/Domain: ankyrin repeat homology <AN13>
F:465-497/Domain: ankyrin repeat homology <AN14>
F:498-530/Domain: ankyrin repeat homology <AN15>
F:531-563/Domain: ankyrin repeat homology <AN16>
F:564-596/Domain: ankyrin repeat homology <AN17>
F:597-629/Domain: ankyrin repeat homology <AN18>
F:630-662/Domain: ankyrin repeat homology <AN19>
F:663-695/Domain: ankyrin repeat homology <AN20>
F:696-728/Domain: ankyrin repeat homology <AN21>
F:729-761/Domain: ankyrin repeat homology <AN22>
F:762-794/Domain: ankyrin repeat homology <AN23>
Query Match 10.4%; Score 457.5; DB 2; Length 1862;
Best Local Similarity 28.7%; Pred. No. 1.3e-16;
Matches 151; Conservative 66; Mismatches 171; Indels 139; Gaps 17;
Qy 60 LNYTE-----NGLSLHLCCICGGKSHIRIMLKGRLPSRLTRNGFTALHLAVYKDN 113
Db 321 LQYNAEIDDIPLHLTPLHVAACGHRV-AKVLDDGAKPNSRALNGFTPLHIAKCKNH 379
Qy 114 AELTSLHSGADTQQVGYGGLTALHTATAGLEAADVLLOGANVNTQDAVFPTPLHI 173
Db 380 IRWELLKKTGASDAVTEGSLPLHVASFMGHLPIVKVLLQKASPNVNVKVTPLHM 439
Qy 174 AAYGHEQVTRLLKFGADVNVSGEDRPLHLASAKGFINTAKLMEFGSKADVNAQDN 233
Db 440 AARAGHTEVAKYLLQNKAKANAKAKDDQTPHCAARIGHTGMVKLLLENG--ASPNLATT 497
Qy 234 EDHVPPLHFCRFGH-----HDIVKY-----LLOSP-- 258
Db 498 AGHTPLHTAAREGHVDFTALLALLEKEASQACMTKGTTPPLHVAAGYKVTALBELLLEHDAH 557
Qy 259 -----LEV-----QHVVNIIYGDTPPLHACYNGKFEVAKEI 289
Db 558 PNAAGKNGTLPPLHVAHHNNLDIVKLLPRGSGPHSPAMNGYTPPLHIAAKQNOIEVARSL 617
Qy 290 IQISGTSLETKENIFSTAFHSACTYKSIDLVKFLDQNV----- 330
Db 618 LQYGG--SANAESVQGVTPPLHLAAQEGHT-EMVALLLSKOANGNKGSLTPPLHLVSOE 674
Qy 331 -----ININHQRDGHTGLHSAACHYGHIRLVQVFLDNGADNMLVACDPSSRS 377
Db 675 GHVLVADVLIKHGVTVDATTMGYTPPLHVAHSGYGNIKVKVFLLOHQADV-----A 725
Qy 378 GEKDEQTCMLWAYEKGHDAIVT--LLKHVKRPQDELPCNEYSPQGGG-----SYVSV 429
Db 726 KTLGLGYSPHQAAQQGHTDITVLLKNGASP-----NEVSSNGTTPPLAIARLGYISVT 779
Qy 430 SPLGKIKSMTEKADILL--RAGLPSPHFH--LQLESEFEFHEIGS 471
Db 780 DVL---KVTIDERSVLVSKHRMSYPTVDEILDVSEDEGDELVS 823

RESULT 10

B35670
protein-tyrosine kinase (EC 2.7.1.112) 2 - slime mold (Dictyostelium discoideum) (frag
C:Species: Dictyostelium discoideum
C:Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 04-Feb-2000
C:Accession: B35670
R:Tan, J.L.; Spudich, J.A.
Mol. Cell. Biol. 10, 3578-3583, 1990
A:Title: Developmentally regulated protein-tyrosine kinase genes in Dictyostelium disc
A:Reference number: A35670; MUID:90287147; PMID:1972546
A:Accession: B35670
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-410 <TAN>
A:Cross-references: GB:M33784; NID:g16777; PIDN:AAA33203.1; PID:g167778
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homo
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F:106-371/Domain: protein kinase homology <KIN>
F:114-122/Region: protein kinase ATP-binding motif
Query Match 10.3%; Score 450.5; DB 2; Length 410;
Best Local Similarity 31.1%; Pred. No. 5e-17;
Matches 131; Conservative 80; Mismatches 145; Indels 65; Gaps 14;
Qy 365 DMNLVACDPSPRSSEGEQTCMLWAYEKGHDAIVT---LLKHVKRPQDELPCNEYSPQGG 421
Db 11 DITFLVCDNPDSSTKEKSNVS-----NTSIIISASNLNRH-----ITPNSHMRPRG 55
Qy 422 DG-SYVSVPSPLGK-----IKSMTEKADILLRAGL-PSHFHLQLSEIFHEIIG 470
Db 56 RSISESLMSPINKESLNDIQRAISEKIKTKFELKSLIGEREYIIDINDIQIKVG 115
Qy 471 SGSGFKYKRCRCKIVAIKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIQVFGACLN 530
Db 116 EGAFSEWEGWKGIIHVAIKLKI-IGDEQPKFEREVQNLKGNHONIVMGICAYK 174
Qy 531 DPSQPAIVTQYISGSLFSLIHE-----QKRILDQLSKLIIVADVAKMEVHLNLTPII 585
Db 175 -PA--CIITEYWAGGSLVNLNPNSSSTPKVKYSPFLVKMATDVALGLLHSIT--IV 229
Qy 586 HRDLNSHNTLLYEDGHVAVDFEGSRFLQSLDEDNNTKQPG---NLRWMAEVTQCTRY 642
Db 230 HRDLTQNTLLDELGNIKISDFGLS--AEKSGESMTWTNGGICNPRWRPELTKNLGHY 287
Qy 643 TIKADVFSALCLWEILTGEIIPFAHLKPAADAAADMAHYHIPPICGYSTPKPISILLRIGW 702
Db 288 SEKVDVYCFSLVWEILTGEIIPFDLQSGQSAQVAYAGLRPPPIPEYCDPELKLTLTQCV 347
Qy 703 NACPEGRPFSEVWVKLECLCN-----TELMSPASSNSSSGSLSP 742
Db 348 EADPNDRPPTIYVVKLEISWNPIGVSDQFYQYSEPTPRLALSQSSNSISLSP 407
Qy 743 S 743
Db 408 T 408
RESULT 11
T01451
protein kinase homolog F2401.i3 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
C:Accession: T01451
R:Shina, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; C
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, January 1998
A:Description: Genomic sequence for Arabidopsis thaliana BAC F2401.
A:Reference number: Z14211
A:Accession: T01451
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-390 <SHI>
A:Cross-references: EMBL:AC003113; NID:g2689438; PID:g2781357; GSPDB:GN00059; ATSP:F24
C:Genetics:

A;Gene: ATSP.F2401.13
A;Map position: 1
A;Introns: 149/3; 301/3
C;Superfamily: Kinase-related transforming protein; protein kinase homology

Query Match 10.1%; Score 444.5; DB 2; Length 390;
Best Local Similarity 31.6%; Pred. No. 9.8e-17;
Matches 120; Conservative 74; Mismatches 129; Indels 57; Gaps 13;

QY 411 LPCNEYSQGGGSGSVSPPLGKTKSMTEKAD-----ILLRA-----GLPSH 455
DB 28 LFCNPSS-----SKTNRYAEAEETMEKFRDMSMESMILESENVENTWEASKGREE 78

QY 456 FHLQLSEIEFBIIGSGSGKGVKRCNKIVAIKRYRANTYCKSDVDM---PCREVISI 512
DB 79 WTADLSQLEIGNKFPASGAHSRIYRIYQKQAVAVKVPRIPTHKEKTRAKLQKQKSEVAL 138

QY 513 LCQLNHPCVIQVGACLNDPSPQFAIVQYISGGSLFSLHBEQXRI-LDLOSLLIIVDVA 571
DB 139 LGRLEPHNIVQFIAAC-KKPPVYCIITEYMSQGNLRMLNKKPYSLSIETVLEALDIS 197

QY 572 KGWELVNLNTOPIIHRDLNSHNLIVEDGHAVVADFGESRFLQSLDNDMTKQPGNL--- 628
DB 198 RGMEXYDH--SQGVHRLDKNNLLNDEMVRVKVADFGISCL-----ETQCREAKGNMGTY 250

QY 629 RWMAPVFTQCTRYTIKADVFYSALCLMEILTGEIIPFAHLKPAAAAADMAVHHIRPPIGY 688
DB 251 RWMAPEMIKE-KPYTRKVDVYSGIVLMEITALLPFPQGMTPVQAAFAVAEKNRPPLPA 309

QY 689 STPKPISSLLIRGNACPGREPERSEVWMLR---ECLC-NIELMSPASSNSSSL----- 740
DB 310 SCQPALAHLIKRCWSENPSKPDFGNIVAVLEKYDECYKGLPLTSHASLTKTKKAILDH 369

QY 741 -----SPSSSSDCLVN 751
DB 370 LKGCVTISISPPSSSSVPVN 389

RESULT 12
S37771
Ankyrin, erythrocyte - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-May-1994 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C;Accession: S37771
R;Birkenmeier, C.S.; White, R.A.; Peters, L.L.; Hall, E.J.; Lux, S.E.; Barker, J.E.
J. Biol. Chem. 268, 9533-9540, 1993
A;Title: Complex patterns of sequence variation and multiple 5' and 3' ends are found an
A;Reference number: S37771; MUID:93252825; PMID:8486643
A;Accession: S37771
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1848 <BIR>
A;Cross-references: EMBL:X69063; NID:g311816; PIDN:CAA48801.1; PID:g311817
C;Superfamily: ankyrin; ankyrin repeat homology
C;Keywords: alternative splicing
F;48-80/Domain: ankyrin repeat homology <AN01>
F;81-113/Domain: ankyrin repeat homology <AN02>
F;114-146/Domain: ankyrin repeat homology <AN03>
F;147-175/Domain: ankyrin repeat homology <AN04>
F;176-208/Domain: ankyrin repeat homology <AN05>
F;209-241/Domain: ankyrin repeat homology <AN06>
F;242-274/Domain: ankyrin repeat homology <AN07>
F;275-307/Domain: ankyrin repeat homology <AN08>
F;308-340/Domain: ankyrin repeat homology <AN09>
F;341-373/Domain: ankyrin repeat homology <AN10>
F;374-406/Domain: ankyrin repeat homology <AN11>
F;407-439/Domain: ankyrin repeat homology <AN12>
F;440-472/Domain: ankyrin repeat homology <AN13>
F;473-505/Domain: ankyrin repeat homology <AN14>
F;506-538/Domain: ankyrin repeat homology <AN15>
F;539-571/Domain: ankyrin repeat homology <AN16>
F;572-604/Domain: ankyrin repeat homology <AN17>
F;605-637/Domain: ankyrin repeat homology <AN18>

F;638-670/Domain: ankyrin repeat homology <AN19>
F;671-703/Domain: ankyrin repeat homology <AN20>
F;704-736/Domain: ankyrin repeat homology <AN21>
F;737-769/Domain: ankyrin repeat homology <AN22>
F;770-802/Domain: ankyrin repeat homology <AN23>

Query Match 10.1%; Score 444; DB 2; Length 1848;
Best Local Similarity 28.5%; Pred. No. 6.7e-16;
Matches 142; Conservative 63; Mismatches 159; Indels 134; Gaps 15;

QY 60 LNYRTE-----NGLSLHLCCICGGKKSHTIRLMLKGLRPSRLTRNGFTALHLAVYKDN 113
DB 329 LOYNRAEIDITLDHJUTPLHVAACHGHRV-AKVLDDKGAKNPSALNGFTPLHIACKNH 387

QY 114 ASLITSLHSGADIQQVGYGGTALHIATIAHLEAADVLLQHGAVNIOADVFPTPLHI 173
DB 388 IRVMEILLKTGASIDAVTESGJTPLHVASFWGHLPIVKNLLQRCGASPNVSNVKVETELHM 447

QY 174 AAYGHEQVTRILLKFGADVNVSGVDRPLHLASAKGFLNIAKLMBEESKADVNAQDN 233
DB 448 AARAGHTEVAKYLLQNKAKAKAKADQDTPLHCAARIGHTGMVKLLLENG--ASPNIATT 505

QY 234 EDHVPLHFCRSFGH-----HDIVKY-----LLQSD-- 258
DB 506 AGHTPLHTAAREGHVDLTALLEKEASOACTKGFPLHVAAYKGVRLAEELLEHDAH 565

QY 259 -----LEV-----QPHVNVNYGDTPLHLACYNGKFEVAKEI 289
DB 566 PNAAGKNGLTPLHVAHVHNNLDIVKLLPRGSGSPHSPAWNGYTPPLHIAAKQNOIEVARSL 625

QY 290 IQISGTESLTKENISETAFHSACTYKGSIDLKVKFLDQNV----- 330
DB 626 LOYGG--SANAESVQGVTPHLHAAQEGHT-EMVALLLSKQANGNLGNKSGLTPLHLVSOE 682

QY 331 -----ININHQGRDGHGTGLHSACYHGHIRLVQFLDNGADMNLVACDPSRSS 377
DB 683 GHVPVADVLLKHGVTVDATTRMGYTPHLVASHYGNIKLVKFLQHQADV-----A 733

QY 378 GEKDRQOTCLMAYEKGHDAIVT-LLKHYKRFODELPCNEYSQPGDGG-----SYVSV 429
DB 734 KTKLGYSPHLQAAQOQHTDITVLLKNGASP-----NEVSSNGTTPLAIAKRLGYISVT 787

QY 430 SPLGKIKSMTKEKADILL 447
DB 788 DVL---KVTDTSTVVLV 802

RESULT 13
T13940
Ankyrin - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C;Accession: T13940
R;Dubreuil, R.R.; Yu, J.
Proc. Natl. Acad. Sci. U.S.A. 91, 10285-10289, 1994
A;Title: Ankyrin and beta-spectrin accumulate independently of alpha-spectrin in Drosoph
A;Reference number: Z17820; MUID:95024098; PMID:7937942
A;Accession: T13940
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1549 <DUE>
A;Cross-references: EMBL:L35601; NID:g557083; PID:g557084; PIDN:AAC37208.1
C;Genetics:
A;Cross-references: FlyBase:FBgn0011747

Query Match 10.1%; Score 443; DB 2; Length 1549;
Best Local Similarity 25.2%; Pred. No. 6.1e-16;
Matches 179; Conservative 98; Mismatches 217; Indels 216; Gaps 29;

QY 57 KVNLYRTEINGLSLHLCCICGGKKSHTIRLML-----KGLRPSRL----- 97
DB 389 KANPNARALNGTPLHIAC----KKNRMVLELLIKHGKANGATGESLTPLHVASFMGC 444

Search completed: September 2, 2004, 14:25:02
Job time : 57 secs



This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2004, 23:38:08 ; Search time 5035.83 Seconds
(without alignments)
17938.068 Million cell updates/sec

Title: US-10-626-173-1

Perfect score: 3025

Sequence: 1 gtcgaccacagctcggcc.....aaaaaaaaaaggcgccgc 3025

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	734.6	24.3	876	14	CF551822
2	721	23.8	756	12	BG227974
3	702	23.2	757	12	BG219957
4	682.8	22.6	813	12	BG209217

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
CF551822
LOCUS AGENCOURT 15595443 NIH MGC 183 Homo sapiens cDNA clone
DEFINITION IMAGE:30530169 5', mRNA sequence.
ACCESSION CF551822
VERSION CF551822.1 GI:34888656
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 876)
NTH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

5	677.4	22.4	793	12	BG216613
6	665	22.0	665	13	BX501623
7	618.6	20.4	902	10	BE897696
8	617.8	20.4	856	13	BX502807
9	598.8	19.8	617	12	BG213152
10	591.8	19.6	617	12	BG216471
11	589.8	19.5	785	12	BG215938
12	589.4	19.5	613	12	BG212638
13	588.2	19.4	629	12	BG186716
14	587.4	19.4	638	9	AL599654
15	587.2	19.4	785	12	BG181490
16	587	19.4	616	12	BG182504
17	586.8	19.4	632	12	BG210010
18	586.2	19.4	616	12	BG187210
19	584	19.3	607	12	BG214260
20	583.4	19.3	604	12	BG221417
21	582.6	19.3	619	12	BG215301
22	582.6	19.3	638	12	BG185641
23	582.2	19.2	616	12	BG218348
24	582	19.2	614	12	BG185642
25	577.2	19.1	618	12	BG215300
26	576.8	19.1	604	12	BG186114
27	576.8	19.1	617	12	BG183556
28	575.6	19.0	614	12	BG198554
29	574	19.0	617	12	BG188258
30	572.8	18.9	617	12	BG204728
31	572.8	18.9	634	12	BG215865
32	571.2	18.9	615	12	BG196050
33	570.6	18.9	616	12	BG190428
34	570.4	18.9	620	12	BG213672
35	568.4	18.8	889	12	BG190429
36	565.8	18.7	612	12	BG196530
37	562.8	18.6	605	12	BG216472
38	558.6	18.5	602	12	BG192433
39	558.4	18.5	625	12	BG219958
40	555.4	18.4	603	12	BG187211
41	552.8	18.3	614	12	BG211040
42	547.8	18.1	609	12	BG205918
43	547	18.1	626	12	BG188259
44	540.8	17.9	892	12	BG202691
45	537.8	17.8	611	12	BG210009

CF551822 876 bp mRNA linear EST 22-SEP-2003

http://image.llnl.gov
 Plate: NDAM618 row: o column: 10
 High quality sequence start: 15
 High quality sequence stop: 670.
 Location/Qualifiers
 1..876

FEATURES

source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30530169"
 /lab_host="DH10B-Tona (T1 and T5 phage resistant)"
 /clone_lib="NIH_MGC_183"
 /note="Organ: Pooled muscle (cardiac and skeletal);
 Vector: pCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2:
 NotI; Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 1.7. Library was constructed by Invitrogen."

ORIGIN

Query Match 24.3%; Score 734.6; DB 14; Length 876;
 Best Local Similarity 97.6%; Pred. No. 2.7e-126;
 Matches 830; Conservative 0; Mismatches 10; Indels 10; Gaps 8;

399 ACTTCTCTGCTCAGTGGAGCTGATATACAGCAGGTGGATACGGTGGCCCTCCTGCTGC 458
 Db |||||
 28 ACTTCTCTGCTTC-CAGTGGAGCTGATATACAGCAGGTGGATACGGTGGCCCTCCTGCTGC 86
 Qy |||||
 459 CTCCTATTTGCTACAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTTGCACCATGGA 518
 Db |||||
 87 CTCCTATTTGCTACAATAGCTGGCCACCTAGAGGCTGCTGATGCTGTTTGCACCATGGA 146
 Qy |||||
 519 GCTATGTCATATCAAGTGCAGTCTTTTCTACTCCATGATATGTCAGGCTACTAT 578
 Db |||||
 147 GCTAATGTCATATCAAGTGCAGTCTTTTCTACTCCATGATATGTCAGGCTACTAT 206
 Qy |||||
 579 GGACATGAACAGTAACTCGGCTCTTTTGAAATTTGGTCTGATGTAATGTAAGTGT 638
 Db |||||
 207 GGACATGAACAGTAACTCGGCTCTTTTGAAATTTGGTCTGATGTAATGTAAGTGT 266
 Qy |||||
 639 GAAGTTGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATCTTGAATATGCAAAA 698
 Db |||||
 267 GAAGTTGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATCTTGAATATGCAAAA 326
 Qy |||||
 699 CTCTTGATGGAAGGAGCAGCAGATGTGATGCTCAAGATATGAGACCATGTC 758
 Db |||||
 327 CTCTTGATGGAAGGAGCAGCAGATGTGATGCTCAAGATATGAGACCATGTC 386
 Qy |||||
 759 CCCTCCATTTCTGTTCTCGATTTGGACCATGATATAGTTAAGTATCTGCTGCAAAAGT 818
 Db |||||
 387 CCCTCCATTTCTGTTCTCGATTTGGACCATGATATAGTTAAGTATCTGCTGCAAAAGT 446
 Qy |||||
 819 GATTTGGAAGTCAACCTCATGTTTAATATCTATGGAGATACCCCTTACACCTGGCA 878
 Db |||||
 447 GATTTGGAAGTCAACCTCATGTTTGAATATCTATGGAGATACCCCTTACACCTGGCA 506
 Qy |||||
 879 TGCTACATGTCAAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGAAACAGAAAGT 938
 Db |||||
 507 TGCTACATGTCAAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGAAACAGAAAGT 566
 Qy |||||
 939 CTGACTAAGGAAACATCTTTCAGTGAAACAGCTTTTTCATAGTCTGTTGTAACCTATGCGAAG 998
 Db |||||
 567 CTGACTAAGGAAACATCTTTCAGTGAAACAGCTTTTTCATAGTCTGTTGTAACCTATGCGAAG 626
 Qy |||||
 999 AGCATTCACCTAGTCAAAATTTCTTTGATCAGAAATGTCTATAAATCAATCAACCAACAGG- 1057
 Db |||||
 627 AGCATTCACCTAGTCAAAATTTCTTTGATCAGAAATGTCTATAAATCAACCAACAGGAG 686
 Qy |||||
 1058 AAGGGATGGGACACT-GGATTAACCTCT-GCTTGCTACCAAGGTACCAATTCG-CCTGGT 1114
 Db |||||
 687 AAGGGATGGGACACTINGGATTAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746
 Qy |||||
 1115 TCAGTTCTTACTGATATGGAGCTGATATGATCTAGT-GGCTTGTGATCCACGAG- 1172
 Db |||||

Db 747 TCAGTTCTTACTGATATGGAGCTGATATGAATCTAGTGGGCTTGTGATCCACGACGN 806
 Qy 1173 TCTAGTGGTGAAGAAGATGACGACACATGTTT--GATGGGCTTATGAAAAAGGCAT 1229
 Db |||||
 807 TCTAGTGGTGAAGAAGATGACGACACATGTTTGGTGGGCTTATGAAAAAGGCAT 866
 Qy 1230 GATGCCATTG 1239
 Db |||||
 867 GATGCCATTG 876

RESULT 2

BG227974
 LOCUS BG227974 756 bp mRNA linear EST 21-APR-2001
 DEFINITION R5741894 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
 ACCESSION BG227974
 VERSION BG227974.1 GI:13748097
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS
 Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
 Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
 Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
 Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,
 Offenbacher, J., Danzig, J. and Ducar, M.

TITLE
 Creation of genome-wide protein expression libraries using random
 activation of gene expression

JOURNAL
 Nat. Biotechnol. 19 (5), 440-445 (2001)

MEDLINE

PUBMED
 11329013

COMMENT

Contact: Scott J. Cain

Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA

Tel: 216 431 9900

Fax: 216 361 9596

Email: scain@athersys.com

High quality sequence stop: 520.

FEATURES

source

1..756
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="Athersys RAGE Library"
 /notes="See 'Creation of Genome-wide Protein Expression
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is Hti080, since a random activation
 method was used, these sequence tags are not necessarily
 expressed in Hti080 under normal circumstances."

ORIGIN

Query Match 23.8%; Score 721; DB 12; Length 756;
 Best Local Similarity 97.3%; Pred. No. 9.6e-124;
 Matches 733; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

593 AACTCGCCTTCTTTTGAATTTGGTCTGATGTAATTAAGTGGTGAAGTTGGAGATAG 652
 Db |||||
 1 AACTCGCCTTCTTTTGAATTTGGTCTGATGTAATTAAGTGGTGAAGTTGGAGATAG 60
 Qy |||||
 653 ACCCTCCACCTAGCATCTGCAAAAGGATCTTGAATATTCGAAACCTTTGATGGAAGA 712
 Db |||||
 61 ACCCTCCACCTAGCATCTGCAAAAGGATCTTGAATATTCGAAACCTTTGATGGAAGA 120
 Qy |||||
 713 AGCAGCAAGCAGATGGAATGCTCAAGATAATGAAGACCATGTGCCACTCCATTTCTG 772
 Db |||||
 121 AGCAGCAAGCAGATGGAATGCTCAAGATAATGAAGACCATGTGCCACTCCATTTCTG 180
 Qy |||||
 773 TTCTCGATTTGACACCATGATATAGTATCTGCTGCAAGATGATTTGGAAGTTCA 832
 Db |||||


```

KEYWORDS      EST.
SOURCE         Homo sapiens (human)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 813)
AUTHORS       Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
               Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
               Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
               Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
               Offenbacher,J., Danzig,J. and Ducar,M.
TITLE         Creation of genome-wide protein expression libraries using random
               activation of gene expression
JOURNAL       Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE       21227151
PUBMED        11329013
COMMENT       Contact: Scott J. Cain
               Athersys, Inc.
               3201 Carnegie Ave, Cleveland, OH 44115, USA
               Tel: 216 431 9900
               Fax: 216 361 9596
               Email: scain@athersys.com
               High quality sequence stop: 427.
               Location/Qualifiers
                   1..813
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /cell_line="HT1080"
                   /clone_lib="Athersys RAGE Library"
                   /note="See 'Creation of Genome-wide Protein Expression
                   Libraries using Random Activation of Gene Expression',
                   Nature Biotechnology, in press. Note that even though the
                   cell type indicated is HT1080, since a random activation
                   method was used, these sequence tags are not necessarily
                   expressed in HT1080 under normal circumstances."

ORIGIN
Query Match      22.6%; Score 682.8; DB 12; Length 813;
Best Local Similarity 95.0%; Pred. No. 1.1e-116;
Matches 780; Conservative 0; Mismatches 33; Indels 8; Gaps 7;

QY 593 AACTCGCCTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTGGAGATAG 652
DB 1 AACTCGCCTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTGAAGTGGAGATAG 60
QY 653 ACCCTCCACTAGCATCTGCAAAAGGATTTCTGATATTTGCAAACTCTTGATGGAAGA 712
DB 61 ACCCTCCACTAGCATCTGCAAAAGGATTTCTGATATTTGCAAACTCTTGATGGAAGA 120
QY 713 AGGCAGCAAGCAGATGTAATGCTCAAGATATGAAGACCATGTCCTCCACTCCATTTCTG 772
DB 121 AGGCAGCAAGCAGATGTAATGCTCAAGATATGAAGACCATGTCCTCCACTCCATTTCTG 180
QY 773 TTCTCGATTGGACACCATGATAGTTAAGTATCTGCTCCTCAAGTGAATTTGGAAGTTCA 832
DB 181 TTCTCGATTGGACACCATGATAGTTAAGTATCTGCTCCTCAAGTGAATTTGGAAGTTCA 240
QY 833 ACCTCATGTTGTTAATCTATGATGAGATACCCCTTACATCGCATGCTGCTACATGGCAA 892
DB 241 ACCTCATGTTGTTAATCTATGATGAGATACCCCTTACATCGCATGCTGCTACATGGCAA 300
QY 893 ATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGATGAAGAAA 952
DB 301 ATTTGAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGATGAAGAAA 360
QY 953 CATCTTCAGTGAACAGCTTTTCATAGTCTGTACCTATGCGAAGAGCATTCACCTAGT 1012
DB 361 CATCTTCAGTGAACAGCTTTTCATAGTCTGTACCTATGCGAAGAGCATTCACCTAGT 420
QY 1013 CAAATTTCTTCTTGATCAGATGTCTATAACATCAACCAAGGAAGGATGGGCACAC 1072
DB 421 CAAATTTCTTCTTGATCAGATGTCTATAACATCAACCAAGGAAGGATGGGCACAC 480

```

```

QY 1073 TGGATTACACTCTGCTTGTCTACACGGTCACATTCGCCTGGTTCAGTTCTTACTGTGATAA 1132
DB 481 TGGATTACACTCTGCTTGTCTACACGGTCACATTCGCCTGGTTCAGTTCTTACTGTGATAA 540
QY 1133 TGGAGCTGATATGAATCTAGTGGCTTGTGATCCACAGCAGTCTAGTGGTGAAGAAAGATGA 1192
DB 541 TGGAGCTGATATGAATCTAGTGGCTTGTGATCCACAGCAGTCTAGTGGTGAAGAA-ATGA 599
QY 1193 GCAGACATGTTTGTGATCGGCTTATGAAAAAGGGCATGATGCCATTGTCACTCTCTGAA 1252
DB 600 GCAGAC-TGTTTGTGATCGGCTTATGAAAAAGGGCATGATGCCATTGTCACTCTCTG-A 658
QY 1253 GCATTATAAGAGACCAACAGATGAATGGCCCTGTATGAATATTCCTCAGCCTGGAGGAGA 1312
DB 659 GCATTATAAGAAAAACCAAGATGAA-TGGCCCTGTGTANTGAAT-TTTTCAGCCTGGAGGAGA 716
QY 1313 TGGCTCCTATGTCTGTCTTCATCACCCCTTGGGGAAGATTAAAGCATGACAAAAGAGAA 1372
DB 717 AGG-TCCTATGTCTGTGGTCC-TCACCCCTTGGGAAGATTAA--GCTGACCAAGAAGA 772
QY 1373 GGCAGATATTCTCTCTCCTTAAGAGCTGGATTGCTTCACATT 1413
DB 773 GCGCAATATCTCTCCAAGAGCGGATTGCTTACATTCCATT 813

RESULT 5
BG215613 LOCUS. 793 bp mRNA linear EST 21-APR-2001
DEFINITION RST36305 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG215613
VERSION BG215613.1 GI:13742634
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
               Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
               Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
               Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
               Offenbacher,J., Danzig,J. and Ducar,M.
TITLE         Creation of genome-wide protein expression libraries using random
               activation of gene expression
JOURNAL       Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE       21227151
PUBMED        11329013
COMMENT       Contact: Scott J. Cain
               Athersys, Inc.
               3201 Carnegie Ave, Cleveland, OH 44115, USA
               Tel: 216 431 9900
               Fax: 216 361 9596
               Email: scain@athersys.com
               High quality sequence stop: 511.
               Location/Qualifiers
                   1..793
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /cell_line="HT1080"
                   /clone_lib="Athersys RAGE Library"
                   /note="See 'Creation of Genome-wide Protein Expression
                   Libraries using Random Activation of Gene Expression',
                   Nature Biotechnology, in press. Note that even though the
                   cell type indicated is HT1080, since a random activation
                   method was used, these sequence tags are not necessarily
                   expressed in HT1080 under normal circumstances."

ORIGIN
Query Match      22.4%; Score 677.4; DB 12; Length 793;
Best Local Similarity 96.4%; Pred. No. 1.2e-115;
Matches 767; Conservative 0; Mismatches 22; Indels 7; Gaps 7;

```

QY	593	AACTGGCCTTCTTTTGA	AAATTGGTGTGATGTAATATGTAAGTGGTGAAGTTGGAGATAG	652
Db	1	AACTGGCCTTCTTTTGA	AAATTGGTGTGATGTAATATGTAAGTGGTGAAGTTGGAGATAG	60
QY	653	ACCCCTCCACCTAGCAT	CTGCAAAAGGATCTTGAATATTGCAAAACTCTTCATGGAAGA	712
Db	61	ACCCCTCCACCTAGCAT	CTGCAAAAGGATCTTGAATATTGCAAAACTCTTCATGGAAGA	120
QY	713	AGGAGCAAAAGCAGAT	GTGAATGCTCAAGATATAGAGACCATGTCCCACTCCATTTCTG	772
Db	121	AGGAGCAAAAGCAGAT	GTGAATGCTCAAGATATAGAGACCATGTCCCACTCCATTTCTG	180
QY	773	TTCTCGATTTGGACACC	ATGATATAGTTAAAGTATCTGTGCAAGTGAATTTGGAAGTTCA	832
Db	181	TTCTCGATTTGTATACCA	TGATATAGTTAAAGTATCTGTGCAAGTGAATTTGGAAGTTCA	240
QY	833	ACCTCATGTTGTTAATAT	CTATGGAGATACCCCTTACACCTGGCATGTCAAAATGGCAA	892
Db	241	ACCTCATGTTGTTAATAT	CTATGGAGATACCCCTTACACCTGGCATGTCAAAATGGCAA	300
QY	893	ATTTCGAAGTTGCCAAG	GAATAATCTCAAATATCAGGAACAGAAGTCTGACATAAGGAAGA	952
Db	301	ATTTCGAAGTTGCCAAG	GAATAATCTCAAATATCAGGAACAGAAGTCTGACATAAGGAAGA	360
QY	953	CATCTTCAGTGAACAC	AGCTTTTCATAGTGTGTTGACTATGGCAAGACATTCGACCTAGT	1012
Db	361	CATCTTCAGCAACAC	AGCTTTTCATAGTGTGTTGACTATGGCAAGACATTCGACCTAGT	420
QY	1013	CAAAATTTCTTTGATCA	GAATGTCATAAATCATCAACCAACCAAGGAGTGGGCAAC	1072
Db	421	CAAAATTTCTTTGATCA	GAATGTCATAAATCATCAACCAACCAAGGAGTGGGCAAC	480
QY	1073	TGGATTAACACTCTGCT	TGCTACCAAGGTCAATTCGCTGGTTGAGTTCTTACTGGATAA	1132
Db	481	TGGATTAACACTCTGCT	TGCTACCAAGGTCAATTCGCTGGTTGAGTTCTTACTGGATAA	540
QY	1133	TGGAGCTGATATGAAT	CTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAAAGATGA	1192
Db	541	TGGAGCTGATATGAAT	CTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAAGAAAGATGA	600
QY	1193	GCAGACATGTTTGATGT	GGGCTTATGAAAAAGGCAATGATGCCATTTGCACACTCCTCAA	1252
Db	601	GCAGACATGTTTGATGT	GGGCTTATGAAAAAGGCAATGATGCCATTTGCACACTCCTCAA	658
QY	1253	GCATTAAGAGACCAAGA	TGAATTTGCCCTGTGATATGAATATTCATGACCTCGGAGGAGA	1312
Db	659	GC-TTATAAGAGACC-CC	AGATGAATTTCCCTG-CAAGGAGA	714
QY	1313	TGGCTCCTATGTCTGT	TTCCATCACCCCTTTGGGCAAGATTTAAAGCATCACAAAAAGCAA	1372
Db	715	TGGCTCCTATGTCTGT	TTCCATCACCCCTTTGGGCAAGATTTAAAGCATCACAAAAAGCAA	773
QY	1373	GGCAGATATTTCTCTC	1388	
Db	774	GGCCGATTTTCTCTC	789	

RESULT 6
BX501623 665 bp mRNA linear EST 04-SEP-2003
BX501623
LOCUS
DEFINITION DKF2p779A1164.r1.779 (synonym: hnccl) Homo sapiens cDNA clone
DEFINITION DKF2p779A1164.5, mRNA sequence.
ACCESSION BX501623
VERSION BX501623.1 GI:32022118
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 665)
Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.

TITLE	EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
JOURNAL	Unpublished (2003)
COMMENT	Contact: MIPS
	MIPS
	Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
	This is the 5' sequence of the clone insert
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
	Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
	sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
	consortium of the German Genome Project.
	No sl sequence available.
	This clone (DKFZp779A1164) is available at the RZPD in Berlin.
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
	Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES	Location/Qualifiers
source	1..665
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="DKFZp779A1164"
	/tissue_type="liver"
	/dev_stage="fetal"
	/lab_host="DH108"
	/clone_lib="779 (synonym: hmccl)"
	/notes="vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiIB"
ORIGIN	
Query Match	22.0%; Score 665; DB 13; Length 665;
Best Local Similarity	100.0%; Pred.No. 2.5e-113; Gaps 0;
Matches 665; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	573 TACTATGGACATGAACAGGTAACTCGCCTCTCTTTTGAATAATTGGTGTGATGTAAATGTA 632
Db	1 TACTATGGACATGAACAGGTAACTCGCCTCTCTTTTGAATAATTGGTGTGATGTAAATGTA 60
QY	633 AGTGGTAGTTGGAGATAGACCCCTCCACCTAGCATCTGCCAAAGAGATCTTGTGATATT 692
Db	61 AGTGGTAGTTGGAGATAGACCCCTCCACCTAGCATCTGCCAAAGAGATCTTGTGATATT 150
QY	693 GCAAAATCTTGATGGAAGAGCAGACAGATGTGAATGCTCAAGATAATGAAGAC 752
Db	121 GCAAAATCTTGATGGAAGAGCAGACAGATGTGAATGCTCAAGATAATGAAGAC 180
QY	753 CATGTCCGACATCAATTTCGTCTCGAATTGGACACCATGATAGTTAGTATCTGCTG 812
Db	181 CATGTCCCATCTCAATTCGTCTCGAATTGGACACCATGATAGTTAGTATCTGCTG 240
QY	813 CAAAGTGATTGGAAGTTCAACTCATGTTGTTAATATCTATGAGATACCCCTTACAC 872
Db	241 CAAAGTGATTGGAAGTTCAACTCATGTTGTTAATATCTATGAGATACCCCTTACAC 300
QY	873 CTGGCATCTGCTACAAATGGCAAAATTGAAGTTGCCAAGGAAATCATCCAAAATATCAGGAACA 932
Db	301 CTGGCATCTGCTACAAATGGCAAAATTGAAGTTGCCAAGGAAATCATCCAAAATATCAGGAACA 360
QY	933 GAAAGTCTGACTAAGGAAACATCTTTCACTGAAACAGCTTTTCATAGTGTGTTGACCTAT 992
Db	361 GAAAGTCTGACTAAGGAAACATCTTTCACTGAAACAGCTTTTCATAGTGTGTTGACCTAT 420
QY	993 GGCAAGAGCATTCGACCTAGTCAAAATTTCTTTTGATCAGAATGTCAATAACATCAACCAC 1052
Db	421 GGCAAGAGCATTCGACCTAGTCAAAATTTCTTTTGATCAGAATGTCAATAACATCAACCAC 480
QY	1053 CAAGGAGGATGGGCACACTGGATACACTCTGCTGCTACCAAGCTGCATTCGCTG 1112
Db	481 CAAGGAGGATGGGCACACTGGATACACTCTGCTGCTACCAAGCTGCATTCGCTG 540
QY	1113 GTTCAGTCTTACTTGGATATGGAGCTGATATGAATCTAGTGGCTTGTGATCCACGAGG 1172
Db	541 GTTCAGTCTTACTTGGATATGGAGCTGATATGAATCTAGTGGCTTGTGATCCACGAGG 600
QY	1173 TCTAGTGGTGAAAGAAGATGAGCAGACATGTTTGATGTGGGCTTATGAAAAAGGGCATGAT 1232

```

Db      601 TCTAGTGGTGAAGAAGATGAGCAGACATGTTTCATGTGGGCTTATGAAAAAGGGCATGAT 660
QY      1233 GCCAT 1237
        |||||
Db      661 GCCAT 665

RESULT 7
BE997696
LOCUS   601439226F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3924055 5',
DEFINITION
        mRNA sequence.
ACCESSION
BE997696
VERSION
BE997696.1 GI:10363419
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 902)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9761 row: d column: 08
High quality sequence stop: 689.

FEATURES
source
1..902
Location/Qualifiers
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:3924055"
    /tissue type="melanotic melanoma"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NIH_MGC_72"
    /note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
    Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
    Average insert size 2 kb. Library constructed by Life
    Technologies."

ORIGIN
Query Match      20.4%; Score 618.6; DB 10; Length 902;
Best Local Similarity 96.2%; Pred. No. 9e-105;
Matches 678; Conservative 0; Mismatches 19; Indels 8; Gaps 4;

QY      379 AGGATAATGCAGATTGATCACTTCTCTGTTTCACAGTGAGCTGATATACAGAGTTG 438
Db      44  AGGATAATGCAGATTGATCACTTCTCTGTTTCACAGTGAGCTGATATACAGAGTTG 103

QY      439 GATACGGTGGCTCAGTCCCTCATATTGCTACATAGCTGGCCACCTAGAGCTGCTG 498
Db      104 GATACGGTGGCTCAGTCCCTCATATTGCTACATAGCTGGCCACCTAGAGCTGCTG 163

QY      499 ATGTCCTGTTGCAACATGGAGCTAATGTCAATATTCAAGATGCAGTTTTTTTTCACATCCAT 558
Db      164 ATGTCCTGTTGCAACATGGAGCTAATGTCAATATTCAAGATGCAGTTTTTTTTCACATCCAT 223

QY      559 TGCATATTGCACGGTACTATGGAATGAACAGGTAATCGCCCTCTCTTTTGAATTTGGTG 618
Db      224 TGCATATTGCACGGTACTATGGAATGAACAGGTAATCGCCCTCTCTTTTGAATTTGGTG 283

QY      619 CTGATGTAAGTGAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAG 678
Db      284 CTGATGTAAGTGAAGTGGTGAAGTTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAG 343

```

```

QY      679 GATTCTTGAATATTGCAAAACTCTTTGATGGAAGAAGGAGCAAGACAGATGTGAATGTC 738
Db      344 GATTCTTGAATATTGCAAAACTCTTTGATGGAAGAAGGAGCAAGACAGATGTGAATGTC 403

QY      739 AAGATAATGAAGACCATGTCCCACTCCATTTCTGTTCTCGATTTCGACACCATGATATAG 798
Db      404 AAGATAATGAAGACCATGTCCCACTCCATTTCTGTTCTCGATTTCGACACCATGATATAG 463

QY      799 TTAAGTATCTGCTGCAAAAGTGATTTCGAAAGTTTCAACCTCATGTTCTTAAATATCTATGAG 858
Db      464 TTAAGTATCTGCTGCAAAAGTGATTTCGAAAGTTTCAACCTCATGTTCTTAAATATCTATGAG 523

QY      859 ATACCCCTTTACACCTGGCATGCTACAATGGCAAAATTTGAAGTTCCCAAGAAATCATCC 918
Db      524 ATACCCCTTTA-ACCTGGCATGCTACAATGGCAAAATTTGAAGTTCCCAAGAAATCATCC 582

QY      919 AATATACAGGAACAGAAAGTCTGACTTAAGGAAACATCTTCAGTGAACACAGCTTTTCATA 978
Db      583 AATATACAGGAACAGAGCTCTGACTTAAGGAAACATCTTCAGTGAACACAGCTTTTCATA 641

QY      979 GTGTTGTACTATGGCAAGAGCATTGACCTAGTCAAAATTTTCTTTCATCAGAATGTCA 1038
Db      642 GTGTTGTACTATGGCAAGAGCATTGAC--TAGTCAATTTCTTTCATCAGAATGTCT 698

QY      1039 TAAACATCAACACCAAGGAAGGATGGGCACACTGGATTACACT 1083
Db      699 TAGACATTCAACACCAAGGAA--GGATGGCACACTGGATTACTCT 740

RESULT 8
BX502807
LOCUS   DXFZp79C2070_r1 779 (synonym: hnccl) Homo sapiens cDNA clone
DEFINITION
        DXFZp79C2070 5', mRNA sequence.
ACCESSION
BX502807
VERSION
BX502807.1 GI:32024614
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 656)
AUTHORS
Bahr,A., Lauber,J., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobbo,G., Han,M. and Wiemann,S.
EST (Bahr,A., Lauber,J., Mewes,H.W., Weil,B., et al.)
JOURNAL
Unpublished (2003)
COMMENT
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DXFZp79C2070) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..656
Location/Qualifiers
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="DKFZp79C2070"
    /tissue type="liver"
    /dev_stage="fetal"
    /lab_host="DH10B"
    /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB"

ORIGIN
Query Match      20.4%; Score 617.8; DB 13; Length 656;
Best Local Similarity 99.5%; Pred. No. 1.4e-104;

```

Matches 619; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 16 CGGCCCTGGAGAGGAAGAACTTATAATAATGGGAAATTTATAAATCTAGACCAACC 75
 |
 Db 35 CTGCCCTGGAGAGGAAGAACTTATAATAATGGGAAATTTATAAATCTAGACCAACC 94
 |

Qy 76 ABACTTGTCTAGTGAATGGAGAGAGAGAGTCAAGTCAATCATATCTTATCAATAGAAA 135
 |
 Db 95 ABACTTGTCTAGTGAATGGAGAGAGAGAGTCAAGTCAATCATATCTTATCAATAGAAA 154
 |

Qy 136 GATTAGAAGATGACCTGCAGATCAAGAGAGAGAGAGTCAAGTCAATCATATCTTATCAATAGAAA 195
 |
 Db 155 GATTAGAAGATGACCTGCAGATCAAGAGAGAGAGAGTCAAGTCAATCATATCTTATCAATAGAAA 214
 |

Qy 196 GCTCTGATGAGCCTTCAGTAAAGTCAATTTAAATACCGCACTGAAATGGCTGCTC 255
 |
 Db 215 GCTCTGATGAGCCTTCAGTAAAGTCAATTTAAATACCGCACTGAAATGGCTGCTC 274
 |

Qy 256 TACTTCATTTATCTGTCATTTGTGGAGGCAAGAAATACATATTCGAACTCTTATGTTGA 315
 |
 Db 275 TACTTCATTTATCTGTCATTTGTGGAGGCAAGAAATACATATTCGAACTCTTATGTTGA 334
 |

Qy 316 AAGGGCTCGCCCATCTCGACTGACAGAGAAATGGAATTAACAGCCTTGCAATTTAGCAGTTT 375
 |
 Db 335 AAGGGCTCGCCCATCTCGACTGACAGAGAAATGGAATTTAGCAGCCTTGCAATTTAGCAGTTT 394
 |

Qy 376 ACAAGGATATGAGAAATGATCACTTCTCTGCTTCCACAGTGGAGCTGATATACAGCAGG 435
 |
 Db 395 ACAAGGATATGAGAAATGATCACTTCTCTGCTTCCACAGTGGAGCTGATATACAGCAGG 454
 |

Qy 436 TTGGATACGGTGGCCCTCACTGCGCCCTCCATATTTGCTACAAATAGCTGCGCACTAGAGGCTG 495
 |
 Db 455 TTGGATACGGTGGCCCTCACTGCGCCCTCCATATTTGCTACAAATAGCTGCGCACTAGAGGCTG 514
 |

Qy 496 CTGATGTCGTTGGCAACATGAGCTGAGTCAATTTCAATTTCAAGATCAGTTTTCATC 555
 |
 Db 515 CTGATGTCGTTGGCAACATGAGCTGAGTCAATTTCAATTTCAAGATCAGTTTTCATC 574
 |

Qy 556 CATTGCATATTTGAGGCTACTATGAGATGAGCAAGTAACTGCGCTTCTTTTGAATTTG 615
 |
 Db 575 CATTGCATATTTGAGGCTACTATGAGATGAGCAAGTAACTGCGCTTCTTTTGAATTTG 634
 |

Qy 616 GTGCTGATGTAATGTAAGTGG 637
 |
 Db 635 GTGCTGATGTAATGTAAGTGG 656
 |

RESULT 9
 BG213152
 LOCUS
 DEFINITION RST32757 Athersys RAGE Library Homo sapiens cdNA, mRNA sequence.
 ACCESSION BG213152
 VERSION BG213152.1 GI:13734839
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 617)
 Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.
 Creation of genome-wide protein expression libraries using random activation of gene expression

TITLE
 JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
 MEDLINE 21227151
 PUBMED 11329013

COMMENT
 Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900

Fax: 216 361 9596
 Email: scain@athersys.com
 High quality sequence stop: 432.
 Location/Qualifiers
 source
 1..617
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="Hri1080"
 /clone_lib="Athersys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is Hri1080, since a random activation method was used, these sequence tags are not necessarily expressed in Hri1080 under normal circumstances."

ORIGIN

Query Match 19.8%; Score 598.8; DB 12; Length 617;
 Best Local Similarity 99.2%; Pred. No. 4.9e-101;
 Matches 611; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 2398 GTGCTGAGCAATATTCCTCTCAAGTCTGTCTTTGAGGAGATGAAAGAGTCTTCAAT 2457
 |
 Db 3 GTGCTGAGCAATATTCCTCTC-AGGTCTGTCTTTGAGGAGATGAAAGAGTCTTCAAT 61
 |

Qy 2458 ACACACCCATTGACAAATATGCTATGATCCGATCCCATGAGCTCAATGCAATTTCAAT 2517
 |
 Db 62 ACACACCCATTGACAAATATGCTATGATCCGATCCCATGAGCTCAATGCAATTTCAAT 121
 |

Qy 2518 CTTGCGCAATATGATGAGCTTTGAGGACAGCAGCTGACAGATTGGCGGTATACCTAAG 2577
 |
 Db 122 CTTGCGCAATATGATGAGCAGCTTTGAGGACAGCAGCTGACAGATTGGCGGTATACCTAAG 181
 |

Qy 2578 GAGAGTTTTTCCCGAAGCTGACAGCAAGATTCCCAACACGCGCAAGCTGCCAAT 2637
 |
 Db 182 GAGAGTTTTTCCCGAAGCTGACAGCAAGATTCCCAACACGCGCAAGCTGCCAAT 241
 |

Qy 2638 ATAAACATTTTACTCTCAAAAGTCTCCTTTAAATTTGGGCTTTTCTTTTACTTGTCTATTTAA 2697
 |
 Db 242 ATAAACATTTTACTCTCAAAAGTCTCCTTTAAATTTGGGCTTTTCTTTTACTTGTCTATTTAA 301
 |

Qy 2698 TTCCCACTATTAGCAGGCTTTGGATTTGCTGCTAAGGAAATATATGCAAGAACCAAG 2757
 |
 Db 302 TTCCCACTATTAGCAGGCTTTGGATTTGCTGCTAAGGAAATATATGCAAGAACCAAG 361
 |

Qy 2758 ACAGATGTATATGAAGAATTTGTTTAAATTTGTAAATTTAAATAAATAATTTAGATCGTT 2817
 |
 Db 362 ACAGATGTATATGAAGAATTTGTTTAAATTTGTAAATTTAAATAAATAATTTAGATCGTT 421
 |

Qy 2818 ACTTGAATGAGAGCTAAAGTCTGTGGGACAGATAATATATATGTTTCTGGGCTGA 2877
 |
 Db 422 ACTTGAATGAGAGCTAAAGTCTGTGGGACAGATAATATATATGTTTCTGGGCTGA 481
 |

Qy 2878 ATTATGTAGACTTTGGTTTGACAGCTATGGGTTTATTTCTTAGAACATTTGTTCTTTCT 2937
 |
 Db 482 ATTATGTAGACTTTGGTTTGACAGCTATGGGTTTATTTCTTAGAACATTTGTTCTTTCT 541
 |

Qy 2938 TTTCTCATTTATGTTACTTCTTAGTGTTCACCTCTGCAATTAAGATTTCTTGTGGAATAG 2997
 |
 Db 542 TTTCTCATTTATGTTACTTCTTAGTGTTCACCTCTGCAATTAAGATTTCTTGTGGAATAG 601
 |

Qy 2998 AAAAAAAAAAAAAAAAAA 3013
 |
 Db 602 AAAAAAAAAAAAAAAAAA 617
 |

RESULT 10
 BG216471
 LOCUS
 DEFINITION RST36050 Athersys RAGE Library Homo sapiens cdNA, mRNA sequence.
 ACCESSION BG216471
 VERSION BG216471.1 GI:13742380

```

KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 617)
AUTHORS    Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
            Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
            Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
            Offenbacher,J., Danzig,J. and Ducar,M.
TITLE      Creation of genome-wide protein expression libraries using random
            activation of gene expression
JOURNAL    Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE    21227151
PUBMED     11329013
COMMENT    Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 371.
FEATURES   source
            1..617
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /cell_line="HT1080"
            /clone_lib="Athersys RAGE Library"
            /note="See 'Creation of Genome-wide Protein Expression
            Libraries using Random Activation of Gene Expression', the
            Nature Biotechnology, in press. Note that even though the
            cell type indicated is HT1080, since a random activation
            method was used, these sequence tags are not necessarily
            expressed in HT1080 under normal circumstances."
ORIGIN
Query Match      19.6%; Score 591.8; DB 12; Length 617;
Best Local Similarity 99.2%; Pred. No. 9.8e-100;
Matches 604; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 2402 TGGCAATATTCCTCTCAAGGTCGTCTTTGGAGGAGATGAAAGAGTCTTCAATACAC 2461
DB 10 TGGCCAAATTCCTCTC-TGGTCGTCTTTGGAGGAGATGAAAGAGTCTTCAATACAC 69
QY 2462 ACCGATTGACAAATATGGCTATGTATCGATCCCATGAGCTCAATGCATTTTCATTCTTG 2521
DB 69 ACCGATTGACAAATATGGCTATGTATCGATCCCATGAGCTCAATGCATTTTCATTCTTG 128
QY 2522 CGGAATAGTAGCAGCTTTGAGGACAGCAGCTGACAGCATTCGGCGTATACCTAAGGAGA 2581
DB 129 CGGAATAGTAGCAGCTTTGAGGACAGCAGCTGACAGCATTCGGCGTATACCTAAGGAGA 188
QY 2582 GTTTTTTCCCGAAGTACGACGACGATTCACACCGCAAGCTGGCTCCAACTATAA 2641
DB 189 GTTTTTTCCCGAAGTACGACGACGATTCACACCGCAAGCTGGCTCCAACTATAA 248
QY 2642 CATTTTACTCTCAAGGTCCTCTTAAATTTGGGCTTTTACTTGTCTTATTTAATTC 2701
DB 249 CATTTTACTCTCAAGGTCCTCTTAAATTTGGGCTTTTACTTGTCTTATTTAATTC 308
QY 2702 CCACATTTAGCAGGCTTTGGATTTGTGCTTAAGGAATAATATGCAAAAGAACCAAGACAG 2761
DB 309 CCACATTTAGCAGGCTTTGGATTTGTGCTTAAGGAATAATATGCAAAAGAACCAAGACAG 368
QY 2762 AATGATATGAGGATTTGTTTTTAATTTCTGTAATTAABAAAATTTAGATCGTTACTT 2821
DB 369 AATGATATGAGGATTTGTTTTTAATTTCTGTAATTAABAAAATTTAGATCGTTACTT 428
QY 2822 GGAAATGGAGCCTAAGTCTGTGTGGGACAGATAATAATATGTTTTCTGGGCTGAATTA 2881
DB 429 GGAAATGGAGCCTAAGTCTGTGTGGGACAGATAATAATATGTTTTCTGGGCTGAATTA 488
QY 2882 TGTAGACTTGTGTTTACAGCTATGGGTTTATTTCTTAGAACAATTGTCATTTCTTTC 2941
DB 489 TGTAGACTTGTGTTTACAGCTATGGGTTTATTTCTTAGAACAATTGTCATTTCTTTC 548
QY 2942 TCATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAAGATCTTTGGTGAATAGAAAA 3001
DB 549 TCATTATGTTACTTCTAGTGTTCACCTCTGTGATTAAAGATCTTTGGTGAATAGAAAA 608
QY 3002 AAAAAA 3010
DB 609 AAAAAA 617
RESULT 11
BG215938
LOCUS      BG215938              785 bp      mRNA      linear      EST 21-APR-2001
DEFINITION RGT35624 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION  BG215938
VERSION    BG215938.1 GI:13741959
KEYWORDS  ESR.
SOURCE     Homo sapiens
ORGANISM  Homo sapiens (human)
REFERENCE  1 (bases 1 to 785)
AUTHORS    Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
            Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
            Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
            Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
            Offenbacher,J., Danzig,J. and Ducar,M.
TITLE      Creation of genome-wide protein expression libraries using random
            activation of gene expression
JOURNAL    Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE    21227151
PUBMED     11329013
COMMENT    Contact: Scott J. Cain
            Athersys, Inc.
            3201 Carnegie Ave, Cleveland, OH 44115, USA
            Tel: 216 431 9900
            Fax: 216 361 9596
            Email: scain@athersys.com
            High quality sequence stop: 550.
FEATURES   source
            1..785
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /cell_line="HT1080"
            /clone_lib="Athersys RAGE Library"
            /note="See 'Creation of Genome-wide Protein Expression
            Libraries using Random Activation of Gene Expression', the
            Nature Biotechnology, in press. Note that even though the
            cell type indicated is HT1080, since a random activation
            method was used, these sequence tags are not necessarily
            expressed in HT1080 under normal circumstances."
ORIGIN
Query Match      19.5%; Score 590.8; DB 12; Length 785;
Best Local Similarity 99.7%; Pred. No. 1.4e-99;
Matches 592; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2420 AGGTCTGCTTTGGAGGAGATGAAAGAGTCTTCAATACACACCCATGCAATATGG 2479
DB 23 AGGTCTGCTTTGGAGGAGATGAAAGAGTCTTCAATACACACCCATGCAATATGG 82
QY 2480 CTATGATCCGATCCCATGAGCTCAATTCATTTTCTTCCGAAATAGTAGCAGCTT 2539
DB 83 CTATGATCCGATCCCATGAGCTCAATTCATTTTCTTCCGAAATAGTAGCAGCTT 142
QY 2540 TGAGGACAGCAGCTGTGACAGCATTCGGCGGTATACCTAAGGAGATTTTTTCCCGAAGTGA 2599
DB 143 TGAGGACAGCAGCTGTGACAGCATTCGGCGGTATACCTAAGGAGATTTTTTCCCGAAGTGA 202

```


EST. Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 638)
Koehler, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.)
EST (Koehler, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp313B0430) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..638
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp313B0430"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="313 (synonym: hlcc2)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match 19.4%; Score 587.4; DB 9; Length 638;
Best Local Similarity 99.8%; Pred. No. 6.4e-99;
Matches 588; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 16 CGGCCCTGGAGAGGAAGAACTTATTAATGGAATTTATAAATCTAGACCAACC 75
Db 50 CTGCCCTGGAGAGGAAGAACTTATTAATGGAATTTATAAATCTAGACCAACC 109
QY 76 AAATCTTCTAGTATGATGGAAGAAAGTTCAGTGAATCATATGTTTCAATAGAAA 135
Db 110 AAATCTTCTAGTATGATGGAAGAAAGTTCAGTGAATCATATGTTTCAATAGAAA 169
QY 136 GATTAGAGATGACCTGCAGATCAAGGAAAAGAACTGACAGAACTAGGATATATTG 195
Db 170 GATTAGAGATGACCTGCAGATCAAGGAAAAGAACTGACAGAACTAGGATATATTG 229
QY 196 GCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTTACCGCACTGAAATGGCTGTCTC 255
Db 230 GCTCTGATGAAGCCTTCAGTAAAGTCAATTTAAATTTACCGCACTGAAATGGCTGTCTC 289
QY 256 TACTTCAATTTATGTTGCAATTTGTGGAGGCAAGAAATCAATATTCGAATCTTATGTTGA 315
Db 290 TACTTCAATTTATGTTGCAATTTGTGGAGGCAAGAAATCAATATTCGAATCTTATGTTGA 349
QY 316 AAGGGCTCCGCCATCTCGACTGACAGAAATGANTTACAGCTTGCATTTAGCAGTTT 375
Db 350 AAGGGCTCCGCCATCTCGACTGACAGAAATGANTTACAGCTTGCATTTAGCAGTTT 409
QY 376 ACAAGGATAATGACAGATTCATCTTCTCTGCTTCAAGTGGAGCTGATATACAGCAGG 435
Db 410 ACAAGGATAATGACAGATTCATCTTCTCTGCTTCAAGTGGAGCTGATATACAGCAGG 469
QY 436 TTGATAGGTTGGCTCCTCCTCCATATTTGCTACATAGCTGCGCACTAGAGGCTG 495
Db 470 TTGATAGGTTGGCTCCTCCTCCATATTTGCTACATAGCTGCGCACTAGAGGCTG 529
QY 496 CTGATGTGCTGTTGCAACATGGAGCTAAATGCTCAATATTCGAATGAGTCTTTCCTC 555

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Email: scain@atersys.com
High quality sequence stop: 629.
Location/Qualifiers
1..629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries Using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

ORIGIN
Query Match 19.4%; Score 588.2; DB 12; Length 629;
Best Local Similarity 97.0%; Pred. No. 4.5e-99;
Matches 610; Conservative 0; Mismatches 18; Indels 1; Gaps 1;
QY 2397 AGTGTGACAAATATCTCTCAGCTCTGCTTGGAGGAGATGAAGAGTCTTCAA 2456
Db 1 AGTGTGACAAATATCTCTCAGCTCTGCTTGGAGGAGATGAAGAGTCTTCAA 59
QY 2457 TACACACCATTTGACAAATATGCTATGATCCGATCCCATGAGCTCAATGCTTTTCA 2516
Db 60 TACACACCATTTGACAAATATGCTATGATCCGATCCCATGAGCTCAATGCTTTTCA 119
QY 2517 TCTTCCGAAATAGTAGCAGCTTTGAGGACAGAGCTGACAGATTCGCGGTATACCTAA 2576
Db 120 TCTTCCGAAATAGTAGCAGCTTTGAGGACAGAGCTGACAGATTCGCGGTATACCTAA 179
QY 2577 GGAGAGTCTTTTCCCGAAGTGCAGCAAGATTTCCACCAAGCTGGCTTCCAAAC 2636
Db 180 GGAGAGTCTTTTCCCGAAGTGCAGCAAGATTTCCACCAAGCTGGCTTCCAAAC 239
QY 2637 TATAACATTTTACTCTCAAGGCTCTCTTAAATGGCTGTTTACTTGCTTCTTATTA 2696
Db 240 TATAACATTTTACTCTCAAGGCTCTCTTAAATGGCTGTTTACTTGCTTCTTATTA 299
QY 2697 ATTCCCGCAATAGTAGCAGCTTTGAGTTCGCTTGAAGTAAATATGCAAGAAACCAA 2756
Db 300 ATTCCCGCAATAGTAGCAGCTTTGAGTTCGCTTGAAGTAAATATGCAAGAAACCAA 359
QY 2757 GACAGATGTATGAGAAATGTTTAAATTTTGTAAATTAATAAATAATTTAGATCGT 2816
Db 360 GACAGATGTATGAGAAATGTTTAAATTTTGTAAATTAATAAATAATTTAGATCGT 419
QY 2817 TACTTGGAAATGGAGCCTTAAGTCTGTGTGACAGATAATAATATGTTTCTGCGGCTG 2876
Db 420 TACTTGGAAATGGAGCCTTAAGTCTGTGTGACAGATAATAATATGTTTCTGCGGCTG 479
QY 2877 AATTATGTAGACTTGTGTTGACAGCTATGGGTTTATTTCTTGAACATTTGTTTCT 2936
Db 480 AATTATGTAGACTTGTGTTGACAGCTATGGGTTTATTTCTTGAACATTTGTTTCT 539
QY 2937 TTTTCTCATTTATGTTTCTTCTAGTGTTCACCTCTGTGATTAAGATTTCTTGTGAATA 2996
Db 540 TTTTCTCATTTATGTTTCTTCTAGTGTTCACCTCTGTGATTAAGATTTCTTGTGAATA 599
QY 2997 GAAAAAAGAAAAAAGGCGCGCGC 3025
Db 600 GAAAAAAGAAAAAAGGCGCGCGC 628

RESULT 14
AL599654
LOCUS
DEFINITION
ACCESSION
VERSION

AL599654 638 bp mRNA linear EST 04-SEP-2003
DKFZp313B0430.r1.313 (synonym: hlcc2) Homo sapiens cDNA clone
DKFZp313B0430.5', mRNA sequence.
AL599654
AL599654.1 GI:15162942

Db 530 CTGATGTCCTGTCGCAACATGAGAGCTAATGCAATATTCAAGATGCAGTGTCTTTTCACTC 589
QY 556 CATTGCATATTGACGCTACTATGACATGACAGGTAACTCGCCTTCT 604
Db 590 CATTGCATATTGACGCTACTATGACATGACAGGTAACTCGCCTTCT 638

RESULT 15
BG181490
LOCUS BG181490 785 bp mRNA linear EST 21-APR-2001
DEFINITION R31338 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG181490
VERSION BG181490.1 GI:13703177
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 785)
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,
Offenbacher, J., Danzig, J. and Ducar, M.,
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
21227151
11329013
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 481.

FEATURES
source
1..785
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/notes="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

ORIGIN
Query Match 19.4%; Score 587.2; DB 12; Length 785;
Best Local Similarity 97.1%; Pred. No. 6.4e-99;
Matches 598; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 2398 GTGCTGGACAATATTCCTCTCAAGGTCCTCTTTGGAGGAGATGAAAGAGTCTTCAAT 2457
Db 4 GTGCTGGACAATATTCCTCTTAAGGTCCTCTTTGGAGGAGATGAAAGAGTCTTCAAT 63
QY 2458 ACACACCCATTGACAAATATGGCTATGATATCGATCCCATGAGCTCAATGCAATTTTCATT 2517
Db 64 ACACACCCATTGACAAATATGGCTATGATATCGATCCCATGAGCTCAATGCAATTTTCATT 123
QY 2518 CTTGCCGAATAGTAGCAGCTTTGAGGACAGAGCTGACAGCATTCGGCGTATACCTTAAG 2577
Db 124 CTTGCCGAATAGTAGCAGCTTTGAGGACAGAGCTGACAGCATTCGGCGTATACCTTAAG 183
QY 2578 GAGAGTTTTTCCCGAATGACAGCAAGCATTCACACCGGCAAGCTGGCTTCCAACT 2637
Db 184 GAGAGTTTTTCCCGAATGACAGCAAGCATTCACACCGGCAAGCTGGCTTCCAACT 243
QY 2638 ATAAACATTTACTCTCAAGGTCCTCTTAAGTGGCTGTGTTTACTGTGCTTATTTAA 2697

Search completed: September 6, 2004, 07:38:00
Job time : 5038.83 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 5, 2004, 19:27:21 ; Search time 631.459 Seconds
(without alignments)
16852.608 Million cell updates/sec

Title: US-10-626-173-3
Perfect score: 2505
Sequence: 1 atgggaattataaaatctag.....gcagctttgagacagcgc 2505

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2505	100.0	2505	3	AAA47607 Human CAR
2	2505	100.0	2508	4	Aaf44702 Novel pro
3	2505	100.0	3025	3	AAA47606 Human CAR
4	2505	100.0	3025	7	ACC48580 Human car
5	2455.4	98.0	2788	5	AAS80720 DNA encod
6	1892.2	75.5	2505	3	AAA47609 Rat CAR
7	1892.2	75.5	3026	3	AAA47608 Rat CAR
8	1892.2	75.5	3026	7	ACC48581 Rat cardi
9	1772	70.7	2024	8	AAD57334 Human kin
10	466.4	18.6	668	6	ABT09088 Phase-1 R
11	458.4	18.3	633	4	AAS33234 DNA encod
12	432	17.2	640	4	AAS33357 DNA encod
13	135.6	5.4	1557	5	AAS80722 DNA encod
14	106.4	4.2	2283	3	AAC48526 Arabidops
15	100.2	4.0	1704	7	ADA69514 Rice gene
16	98	3.9	1662	6	ABZ13209 Arabidops
17	91.4	3.6	3323	9	ADB62159 Human cDN
18	87.8	3.5	1845	7	ADA70011 Rice gene
19	87.4	3.5	1299	4	AAS33288 DNA encod
20	87.4	3.5	1299	5	ABK72043 Human cDN
21	87.4	3.5	1299	6	ABA14324 Human ner
22	87.4	3.5	1299	6	ABK91635 cDNA enco
23	85.6	3.4	3553	9	ADC39121 Novel hum

24	85.6	3.4	4273	4	ABS54855	Human ank
25	85.6	3.4	4519	7	AAD51564	Human str
26	84.6	3.4	1514	3	AAC47638	Arabidops
27	84.2	3.4	5175	6	AAX99407	DNA of AP
28	84	3.4	2915	6	ABN59866	Novel hum
29	84	3.4	13329	5	AAS71562	DNA encod
30	84	3.4	14770	6	ABJ61893	Colon ade
31	84	3.4	14770	6	ABJ68857	Kidney ca
32	84	3.4	14770	6	ABJ68858	Kidney ca
33	84	3.4	14770	9	ADB31314	Testoster
34	84	3.4	14797	5	AAS65159	DNA encod
35	84	3.4	15187	5	AAS86131	DNA encod
36	82.6	3.3	6278	4	AAK52293	Human pcl
37	80.8	3.2	1797	5	AAK50586	Degenerat
38	80.8	3.2	1797	6	AAK36756	Human RIN
39	80.6	3.2	3160	4	AAC90485	Rice EDR1
40	80.2	3.2	883	6	ABK65288	Arabidops
41	80.2	3.2	883	9	ADD30938	Plant yie
42	80.2	3.2	883	9	ADG37244	Plant yie
43	79.6	3.2	3438	9	ADB62405	Human cDN
44	79.4	3.2	1914	7	ADA69560	Rice gene
45	79.4	3.2	2187	7	ADA70219	Rice gene

ALIGNMENTS

RESULT 1
AAA47607
ID AAA47607 standard; cDNA; 2505 BP.
XX
AC AAA47607;
XX
DT 20-OCT-2000 (first entry)
XX
DE Human CAR (Cardiac related Ankyrin-Repeat Protein Kinase) cDNA.

XX
KW Cardiac related ankyrin repeat protein kinase; CAR; cytoskeleton;
KW cardiac cell growth factor receptor; cell differentiation; modulator;
KW regulator; detection; cellular proliferation; cardiovascular disorder;
KW heart failure; hypertension; cancer; sarcoma; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..2505
FT /*tag= a
FT /product= "Human CAR"

PN WO200034330-A1.

XX
XX 15-JUN-2000.
XX
XX 10-DEC-1999; 99WO-US029465.
XX
XX 11-DEC-1998; 98US-0111938P.
XX 14-APR-1999; 99US-00291839.
XX (MILL-) MILLENNIUM PHARM INC.

XX Raju J;
XX WFI; 2000-431275/37.
XX P-PSDB; AAB01470.

XX New polynucleotide encoding cardiac-related ankyrin-repeat protein
XX kinase, useful for treating disorders such as cardiovascular disorders,
XX e.g. heart failure and cell differentiation disorders, e.g. cancer.

XX Claim 1; Page 143-147; 16ipp; English.

XX CARC polypeptides are regulators of signal transmission from cellular
XX receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton

CC function. They also act as modulators of cellular differentiation and
CC cell death. Nucleic acids, or their fragments encoding CRK polypeptides
CC are useful for detecting CRK nucleic acids especially mRNA, in a sample.
CC CRK is useful for treating disorders associated with upregulation or
CC downregulation of cellular proliferation such as, cardiovascular
CC disorders (heart failure and hypertension) and disorders associated with
CC cell differentiation such as cancer and sarcoma
XX

SQ Sequence 2505 BP; 721 A; 532 C; 554 G; 698 T; 0 U; 0 Other;

Query Match	100.0%	Score 2505;	DB 3;	Length 2505;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2505;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	ATGGGAATTAATAATCTAGACCAACCCAACTTGACTGATGAATGGAAGAAAAGTC	60		
DB 1	ATGGGAATTAATAATCTAGACCAACCCAACTTGACTGATGAATGGAAGAAAAGTC	60		
QY 61	AGTGAATCATATGTTATCACAATAGAAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA	120		
DB 61	AGTGAATCATATGTTATCACAATAGAAAAGATTAGAAGATGACCTGCAGATCAAGGAAAA	120		
QY 121	GAACTGCAGAACCTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA	180		
DB 121	GAACTGCAGAACCTAAGGAATATATTTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA	180		
QY 181	AATTACCGCACTGAAAATGGGCTGTCTCTACTTCATTTATGTTGCAATTTGTGAGGCAAG	240		
DB 181	AATTACCGCACTGAAAATGGGCTGTCTCTACTTCATTTATGTTGCAATTTGTGAGGCAAG	240		
QY 241	AAATCACATATTCGAATCTATGTTGAAAGGGCTCGCCCATCTCGACTGACAGAAAT	300		
DB 241	AAATCACATATTCGAATCTATGTTGAAAGGGCTCGCCCATCTCGACTGACAGAAAT	300		
QY 301	GGATTTACAGCTTGCAATTTAGCAGTTTACAAGGATAATGCAGAAATTTGATCACTTCTCTG	360		
DB 301	GGATTTACAGCTTGCAATTTAGCAGTTTACAAGGATAATGCAGAAATTTGATCACTTCTCTG	360		
QY 361	CTTCAAGTGGAGTGATATACAGCAGGTTGGATACGGTGGCTCACTGCGCTCCATATT	420		
DB 361	CTTCAAGTGGAGTGATATACAGCAGGTTGGATACGGTGGCTCACTGCGCTCCATATT	420		
QY 421	GCTACAAATAGCTGGCCACTAGAGGCTGTGATGCTGTGTGCAACATGAGCTAATGTC	480		
DB 421	GCTACAAATAGCTGGCCACTAGAGGCTGTGATGCTGTGTGCAACATGAGCTAATGTC	480		
QY 481	AATATTCAGATGTCAGTTTTTTTCACTCCATTTGCATATTCGACGGTACTATGGACATGAA	540		
DB 481	AATATTCAGATGTCAGTTTTTTTCACTCCATTTGCATATTCGACGGTACTATGGACATGAA	540		
QY 541	CAGGTAACCTCGCTCTTTTGAAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA	600		
DB 541	CAGGTAACCTCGCTCTTTTGAAATTTGGTGTGATGTAATGTAAGTGGTGAAGTTGGA	600		
QY 601	GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATTTGCAAACTCTTTGATG	660		
DB 601	GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTCTTGAATTTGCAAACTCTTTGATG	660		
QY 661	GAAGAAGCAGCAAGCAGATGTAATGCTCAAGATTAATGAAGACCATGTCCTCCATCCAT	720		
DB 661	GAAGAAGCAGCAAGCAGATGTAATGCTCAAGATTAATGAAGACCATGTCCTCCATCCAT	720		
QY 721	TTCTCTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTTGGAA	780		
DB 721	TTCTCTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTTGGAA	780		
QY 781	GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATCTACAT	840		
DB 781	GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTACACCTGGCATCTACAT	840		
QY 841	GGCAAAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG	900		
DB 841	GGCAAAATTTGAAGTTGCCAAGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG	900		

QY 901	GAAGAAATCTTTCAGTGAAACAGCTTTTTCATAGTGTCTGTATGCTATGGAAGAGATTGAC	960
DB 901	GAAGAAATCTTTCAGTGAAACAGCTTTTTCATAGTGTCTGTATGCTATGGAAGAGATTGAC	960
QY 961	CTAGTCMAATTTCTTCTTGATCAGATGTCATAAATCATCAACCAACCAAGGAGGATGGG	1020
DB 961	CTAGTCMAATTTCTTCTTGATCAGATGTCATAAATCATCAACCAACCAAGGAGGATGGG	1020
QY 1021	CACATGGATTACACTCTGCTTGCTACCAAGGCTCACTTCGCTGCTGCTTCTTCTTCTG	1080
DB 1021	CACATGGATTACACTCTGCTTGCTACCAAGGCTCACTTCGCTGCTGCTTCTTCTTCTG	1080
QY 1081	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATGCCAGAGCTCTAGTGGTGAATAA	1140
DB 1081	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATGCCAGAGCTCTAGTGGTGAATAA	1140
QY 1141	GATGACGACATGTTTGTGATGGGCTTATGAAAAGGGCATGATGCCATTTGTCACTC	1200
DB 1141	GATGACGACATGTTTGTGATGGGCTTATGAAAAGGGCATGATGCCATTTGTCACTC	1200
QY 1201	CTGAAGCATTTAAGAGACCAACAAGATGAATTCCTGTGAATGAATTTCTCAGCTGGA	1260
DB 1201	CTGAAGCATTTAAGAGACCAACAAGATGAATTCCTGTGAATGAATTTCTCAGCTGGA	1260
QY 1261	GGAGATGGCTCTTATGCTGTCTTCATCAACCTTGGGGAAGATTAAAAGCATGACAAA	1320
DB 1261	GGAGATGGCTCTTATGCTGTCTTCATCAACCTTGGGGAAGATTAAAAGCATGACAAA	1320
QY 1321	GAGAAGCAGATATTTCTCTCTTAAGAGCTGGAATTCCTTCACTTTCCATTTCCAGTCT	1380
DB 1321	GAGAAGCAGATATTTCTCTCTTAAGAGCTGGAATTCCTTCACTTTCCATTTCCAGTCT	1380
QY 1381	TCAGAAATTCAGTTCCATGAGATTTATGCTCAGGTTCTTTTGGGAGAGTATATAAGGA	1440
DB 1381	TCAGAAATTCAGTTCCATGAGATTTATGCTCAGGTTCTTTTGGGAGAGTATATAAGGA	1440
QY 1441	CGATGCAAGAAATATAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCAG	1500
DB 1441	CGATGCAAGAAATATAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCAG	1500
QY 1501	TCAGATGTGATATGTTTTGCCAGAGGTGTCCATTTCTGTCAGCTCAATCATCCCTGC	1560
DB 1501	TCAGATGTGATATGTTTTGCCAGAGGTGTCCATTTCTGTCAGCTCAATCATCCCTGC	1560
QY 1561	GTAATTCAGTTCTGCTGGTCTTGTGAATGATCCAGCAGCTTTGCCATTTGCTCACTCAA	1620
DB 1561	GTAATTCAGTTCTGCTGGTCTTGTGAATGATCCAGCAGCTTTGCCATTTGCTCACTCAA	1620
QY 1621	TACATATCAGGGGTTCTCTGTTCTCCCTCTTCTATGAGCAGAGGATTTCTTGAATTG	1680
DB 1621	TACATATCAGGGGTTCTCTGTTCTCCCTCTTCTATGAGCAGAGGATTTCTTGAATTG	1680
QY 1681	CAGCTTAATTAATTAATTTGCACTAGATGTTGCCAAGGATGAGTACTTCCACAACTG	1740
DB 1681	CAGCTTAATTAATTAATTTGCACTAGATGTTGCCAAGGATGAGTACTTCCACAACTG	1740
QY 1741	ACACAGCCAAATTAATACATCGTGACTTTGAACAGTCAATATTTCTTCTATGAGGATGG	1800
DB 1741	ACACAGCCAAATTAATACATCGTGACTTTGAACAGTCAATATTTCTTCTATGAGGATGG	1800
QY 1801	CATGCTGTGTGGCAGATTTGGAGAAATCAAGATTTCTACGTTCTCTGGTGAACACAC	1860
DB 1801	CATGCTGTGTGGCAGATTTGGAGAAATCAAGATTTCTACGTTCTCTGGTGAACACAC	1860
QY 1861	ATGCAAAAACAACTCGGAACCTCCCTGATGGCTCTCTGAGTGTTCACGAGTGCAC	1920
DB 1861	ATGCAAAAACAACTCGGAACCTCCCTGATGGCTCTCTGAGTGTTCACGAGTGCAC	1920
QY 1921	CGGTACACCAATCAAGCAGATGTTTCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	1980
DB 1921	CGGTACACCAATCAAGCAGATGTTTCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	1980

PN WO2000034330-A1.
XX 15-JUN-2000.
XX 10-DEC-1999; 99WO-US029465.
XX 11-DEC-1998; 98US-0111938P.
PR 14-APR-1999; 99US-00291839.
XX (MILL-) MILLENNIUM PHARM INC.
XX Raju J;
XX WPI; 2000-431275/37.
DR P-PSDB; AAB01470.
XX New polynucleotide encoding cardiac-related ankyrin-repeat protein
PT kinase, useful for treating disorders such as cardiovascular disorders,
PT e.g. heart failure and cell differentiation disorders, e.g. cancer.
XX Claim 1; Fig 1; 161pp; English.
XX CARP polypeptides are regulators of signal transmission from cellular
CC receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton
CC function. They also act as modulators of cellular differentiation and
CC cell death. Nucleic acids, or their fragments encoding CARP polypeptides
CC are useful for detecting CARP nucleic acids especially mRNA, in a sample.
CC CARP is useful for treating disorders associated with upregulation or
CC downregulation of cellular proliferation such as, cardiovascular
CC disorders (heart failure and hypertension) and disorders associated with
CC cell differentiation such as cancer and sarcoma
XX Sequence 3025 BP; 882 A; 622 C; 655 G; 866 T; 0 U; 0 Other;
Query Match 100.0%; Score 2505; DB 3; Length 3025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGAATTATTAATCTAGACCAACCCAACTTGTCTACTGATGGAATGGAAGAAAAGTC 60
DB 48 ATGGGAATTATTAATCTAGACCAACCCAACTTGTCTACTGATGGAATGGAAGAAAAGTC 107
QY 61 AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGATGACCTGAGATCAAGGAAAAA 120
DB 108 AGTGAATCATATGTTATCAATAGAAAGATTAGAAAGATGACCTGAGATCAAGGAAAAA 167
QY 121 GAATGACAGAACTAAGGAATATATTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 180
DB 168 GAATGACAGAACTAAGGAATATATTGGCTCTGATGAAGCCCTTCAGTAAAGTCAATTTA 227
QY 181 AATTACCGCACTGAAATGGCTGTCTCTACTTCAATTTATGTTGATTTGGAGGCAAG 240
DB 228 AATTACCGCACTGAAATGGCTGTCTCTACTTCAATTTATGTTGATTTGGAGGCAAG 287
QY 241 AATACATATTCGAATCTTATGTTGAAGGGCTCCGCCCATCTCGACTGACAGAAAT 300
DB 288 AATACATATTCGAATCTTATGTTGAAGGGCTCCGCCCATCTCGACTGACAGAAAT 347
QY 301 GGATTTACAGCCTTGCAATTTAGCAGTTTACAGGATAATGCAATTTGATCACTTCTCTG 360
DB 348 GGATTTACAGCCTTGCAATTTAGCAGTTTACAGGATAATGCAATTTGATCACTTCTCTG 407
QY 361 CTTTCACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCAGCTCCCTCCATATT 420
DB 408 CTTTCACAGTGGAGCTGATATACAGAGGTTGGATACGGTGGCTCAGCTCCCTCCATATT 467
QY 421 GCTACAAATAGTGGCCACCTAGAGGCTGTGATGTGCTGTGCTTGCACATGGAGCTTAATGTC 480
DB 468 GCTACAAATAGTGGCCACCTAGAGGCTGTGATGTGCTGTGCTTGCACATGGAGCTTAATGTC 527
QY 481 AATATTCAAGATCAGTCTTTTCTTCTACTTCCATTTGATGAGCGCTTACTATGGACATGAA 540
DB 528 AATATTCAAGATCAGTCTTTTCTTCTACTTCCATTTGATGAGCGCTTACTATGGACATGAA 587

QY 541 CAGGTAACCTCGCTCTCTTTTGAATTTGGTGTGATGTAATGTAAAGTGGTGAAGTTGGA 600
DB 588 CAGGTAACCTCGCTCTCTTTTGAATTTGGTGTGATGTAATGTAAAGTGGTGAAGTTGGA 647
QY 601 GATAGACCCCTCCACCTAGCATCTGCAAAAAGGATTTCTTGAATATATGCAAAACTCTTGATG 660
DB 648 GATAGACCCCTCCACCTAGCATCTGCAAAAAGGATTTCTTGAATATATGCAAAACTCTTGATG 707
QY 661 GAAGAAGGCAGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720
DB 708 GAAGAAGGCAGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 767
QY 721 TTCTGTTCTCGATTTGGACACCATGATATAGTAAATATCTGTGCAAAAGTGAATTTGAA 780
DB 768 TTCTGTTCTCGATTTGGACACCATGATATAGTAAATATCTGTGCAAAAGTGAATTTGAA 827
QY 781 GTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTACACCTGCGATGCTACAAT 840
DB 828 GTTCAACCTCATGTTGTTAAATATCTATGGAGATACCCCTTACACCTGCGATGCTACAAT 887
QY 841 GGCAAAATTTGAAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
DB 888 GGCAAAATTTGAAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 947
QY 901 GAAACATCTTCAGTGAACACGCTTTTCATAGTGTCTGTACTATGCGCAAGCATTTGAC 960
DB 948 GAAACATCTTCAGTGAACACGCTTTTCATAGTGTCTGTACTATGCGCAAGCATTTGAC 1007
QY 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCTATAACATCAACCCACCAAGGAGGATGGG 1020
DB 1008 CTAGTCAAAATTTCTTCTGATCAGAAATGTCTATAACATCAACCCACCAAGGAGGATGGG 1067
QY 1021 CACACTGGATTAACATCTGCTTGTGTACCAAGGTCATTCGCTGCTGTTCACTTCTTACTG 1080
DB 1068 CACACTGGATTAACATCTGCTTGTGTACCAAGGTCATTCGCTGCTGTTCACTTCTTACTG 1127
QY 1081 GATAATGGAGCTGATATGAATCTAGTGGCTGTGTATCCACAGCTCTAGTGGTGAAGAAA 1140
DB 1128 GATAATGGAGCTGATATGAATCTAGTGGCTGTGTATCCACAGCTCTAGTGGTGAAGAAA 1187
QY 1141 GATGACGACATATGTTGATGTTGGCTTTATGAAAAGGGCATGATGCCATTTGTCACTC 1200
DB 1188 GATGACGACATATGTTGATGTTGGCTTTATGAAAAGGGCATGATGCCATTTGTCACTC 1247
QY 1201 CTGAGCATTAAGAGACCAACAGATGAATTCGCTGTGAATGATATTCACGCTGGA 1260
DB 1248 CTGAGCATTAAGAGACCAACAGATGAATTCGCTGTGAATGATATTCACGCTGGA 1307
QY 1261 GGAGATGGCTCTTATGTTCTTCCATCACCTTTGGGGAAGATTAAAGCATGACAAAA 1320
DB 1308 GGAGATGGCTCTTATGTTCTTCCATCACCTTTGGGGAAGATTAAAGCATGACAAAA 1367
QY 1321 GAGAAGGCAGATATTTCTCTCTTAAGAGTGGATTCGCTTCAATTTCCATTTCCAGCTC 1380
DB 1368 GAGAAGGCAGATATTTCTCTCTTAAGAGTGGATTCGCTTCAATTTCCATTTCCAGCTC 1427
QY 1381 TCAGAAATTTGAGTTCCATGAGATTAATGGCTCAGCTCTTTTGGGAAAGTATATAAGGA 1440
DB 1428 TCAGAAATTTGAGTTCCATGAGATTAATGGCTCAGCTCTTTTGGGAAAGTATATAAGGA 1487
QY 1441 CGATGCAAGAAATAAATAGTGGCTATAAAAGTTCATCGAGCCAATACCTACTGCTCAAG 1500
DB 1488 CGATGCAAGAAATAAATAGTGGCTATAAAAGTTCATCGAGCCAATACCTACTGCTCAAG 1547
QY 1501 TCAGATGTGGATATGTTTGGCGAGAGGTGCCATTCCTGCCAGCTCAATCATCCCTGC 1560
DB 1548 TCAGATGTGGATATGTTTGGCGAGAGGTGCCATTCCTGCCAGCTCAATCATCCCTGC 1607
QY 1561 GTAAATTCAGTTTGGGTGCTTGTGTTGAATGATCCAGCCAGTTTGGCAATGTGCTCACTCA 1620
DB 1608 GTAAATTCAGTTTGGGTGCTTGTGTTGAATGATCCAGCCAGTTTGGCAATGTGCTCACTCA 1667

QY 1621 TACATATCAGGGGTTCTCTGTTCTCCCTTCATGACGAGAGGATTCCTGATTG 1680
Db 1668 TACATATCAGGGGTTCTCTGTTCTCCCTTCATGACGAGAGGATTCCTGATTG 1727
QY 1681 CAGTCTAAATTAATTTATGAGTAGATGTTGCCAAGGCAATGAGTACCTTCACACCTG 1740
Db 1728 CAGTCTAAATTAATTTATGAGTAGATGTTGCCAAGGCAATGAGTACCTTCACACCTG 1787
QY 1741 ACACAGCAATATACATCGTGAATGAAACAGTCAATATTTCTCTATGAGATGG 1800
Db 1788 ACACAGCAATATACATCGTGAATGAAACAGTCAATATTTCTCTATGAGATGG 1847
QY 1801 CATCTGTGGGCGAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAC 1860
Db 1848 CATCTGTGGGCGAGATTTGGAGATCAAGATTTCTACAGTCTCTGGATGAAGACAC 1907
QY 1861 ATGACAAACAACTGGGAACCTCGTTGGATGGCTCTGAGGTGTTCAACGAGTGCAT 1920
Db 1908 ATGACAAACAACTGGGAACCTCGTTGGATGGCTCTGAGGTGTTCAACGAGTGCAT 1967
QY 1921 CGGTACACCATCAAGCAGATGCTTCAGCTATGCTCTGTCTGTGGAAATTCCTACT 1980
Db 1968 CGGTACACCATCAAGCAGATGCTTCAGCTATGCTCTGTCTGTGGAAATTCCTACT 2027
QY 1981 GCGCAAAATTCATTCGCTCATCTCAAGCAGCGCTCGGCGACGACATGCTTACCAC 2040
Db 2028 GCGCAAAATTCATTCGCTCATCTCAAGCAGCGCTCGGCGACGACATGCTTACCAC 2087
QY 2041 CACATCAGACTCCCATGGCTATTCATTCACCAAGCCCATATCATCTCTGCTGATACA 2100
Db 2088 CACATCAGACTCCCATGGCTATTCATTCACCAAGCCCATATCATCTCTGCTGATACA 2147
QY 2101 GGTGGACGATGCTCTGAGGAGACCGGAATTTCTCAAGTGTCTCATGAATGAGAA 2160
Db 2148 GGTGGACGATGCTCTGAGGAGACCGGAATTTCTCAAGTGTCTCATGAATGAGAA 2207
QY 2161 GAGTGTCTCTCAACATGAGCTGATGCTCTCTGCAATCAAGTAAACAGCAGTGGTCTTC 2220
Db 2208 GAGTGTCTCTCAACATGAGCTGATGCTCTCTGCAATCAAGTAAACAGCAGTGGTCTTC 2267
QY 2221 TCACCTCTCTCTCTGATGCTGCTGTAACCGGGAGGACCTGGCCGAGTCAATGG 2280
Db 2268 TCACCTCTCTCTCTGATGCTGCTGTAACCGGGAGGACCTGGCCGAGTCAATGG 2327
QY 2281 GCAGATTAAGAAGTCGTTTCGAATTCGAATATGCTCTAAATCAAGAGTCTTATGCTGT 2340
Db 2328 GCAGATTAAGAAGTCGTTTCGAATTCGAATATGCTCTAAATCAAGAGTCTTATGCTGT 2387
QY 2341 TTGTCCCAAGTCTGACATATATTCCTCTCAAGGTCTGTCTTGGAGGATGAAGA 2400
Db 2388 TTGTCCCAAGTCTGACATATATTCCTCTCAAGGTCTGTCTTGGAGGATGAAGA 2447
QY 2401 AGTCTTCAATACACACCATGACAAATATGGTATGATATCCGATCCCATGAGCTCAATG 2460
Db 2448 AGTCTTCAATACACACCATGACAAATATGGTATGATATCCGATCCCATGAGCTCAATG 2507
QY 2461 CATTTTCATTTCCGAAATAGTAGAGTCTTGGAGCAGCAGC 2505
Db 2508 CATTTTCATTTCCGAAATAGTAGAGTCTTGGAGCAGCAGC 2552

RESULT 4
ID ACC48580 standard; cDNA; 3025 BP.
XX
AC ACC48580;
XX
DT 11-AUG-2003 (first entry)
XX
DE Human cardiac-related ankyrin-repeat protein kinase (CARK) cDNA.
XX
KW Human; CARK; cardiac-related ankyrin-repeat protein kinase; enzyme;
KW cardiant; hypotensive; cytostatic; chromosome 1; gene therapy; gene; ss.

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 48..2555
FT /*tag= a
FT /product= "Human CARK"
FT /note= "the CDS is also claimed in Claim 1"
XX
XX MO2003020912-A2.
XX
XX 13-MAR-2003.
XX
XX 04-SEP-2002; 2002WO-US028300.
XX
XX 05-SEP-2001; 2001US-00947199.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Raju J;
XX
XX WPI; 2003-230188/28.
XX
XX Novel isolated cardiac-related ankyrin-repeat protein kinase polypeptide,
XX useful for treating cellular growth related disorders which include
XX cardiovascular disorders and proliferative and/or differentiative
XX disorders.
XX
XX Claim 1; Fig 1A; 158pp; English.

The present sequence, the coding region of which is also claimed, is that of cDNA encoding a novel human cardiac-related ankyrin-repeat protein kinase, designated CARK, which plays a role in signalling pathways associated with cardiac cell growth or differentiation. The cDNA was isolated from a library prepared from subjects suffering from congestive heart failure of ischaemic and idiopathic origin. A plasmid containing the human CARK nucleotide sequence is deposited as ATCC PTA-1530. The CARK gene maps to chromosome 1, within the atrioventricular canal defects (AVCD) locus. The gene is therefore a candidate for a congenital heart defect susceptibility gene. CARK gene expression is increased in ischaemic heart tissue samples, further suggesting a role in the regulation of cardiac cell growth and/or differentiation and the pathogenesis of cardiovascular disorders, e.g. congestive heart failure and cardiac hypertrophy. The invention provides CARK nucleic acids, antisense molecules, recombinant expression vectors, host cells and transgenic animals in which a CARK gene has been introduced or disrupted, and also CARK proteins, fusion proteins, antigenic peptides and anti-CARK antibodies. These are useful as modulating agents for regulating a variety of cellular processes, e.g. cardiac cellular process, for modulating the phosphorylation state of a CARK molecule or one or more proteins involved in cellular growth or differentiation, for modulating cell behaviour or as targets and therapeutic agents controlling cardiac cell proliferation, differentiation, hypertrophy and migration, for modulating intra-or inter-cellular signalling and/or gene transcription, for modulating cell proliferation, growth, differentiation, survival and/or migration, for regulating transmission of signals from cellular receptors, for modulating entry of cells, e.g. cardiac precursor cells, into mitosis, or for regulating cytoskeletal function. The nucleic acids and proteins are useful for treating cellular growth related disorders which include cardiovascular disorders (such as heart failure, hypertension), and proliferative and/or differentiative disorders (such as cancer). They are also useful in screening assays, detection assays (e.g. chromosomal mapping, tissue typing, forensic biology), predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials) and pharmacogenomics, and in methods of treatment. The nucleic acids are also useful as probes and primers, in gene therapy, in the detection of genetic alteration in the CARK gene, and in forensic identification

XX Sequence 3025 BP; 882 A; 622 C; 655 G; 866 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 2505; DB 7; Length 3025;
XX Best Local Similarity 100.0%; Pred. No. 0;

Matches 2505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGGGAAATTTAAATCTAGACCAACCCAAAACCTTGCTGATGATGAATGGAAGAAAAGTC	60
Db	48	ATGGGAAATTTAAATCTAGACCAACCCAAAACCTTGCTGATGATGAATGGAAGAAAAGTC	107
Qy	61	AGTGAATCATATGCTATCAATAGAAAGATTAGAGATGACCTCAGATCAAGGAAA	120
Db	108	AGTGAATCATATGCTATCAATAGAAAGATTAGAGATGACCTCAGATCAAGGAAA	167
Qy	121	GAATGACAGAACTAAGGAATATTTGGCTCTGATGAAGCCCTTCACTAAAGTCAATTTA	180
Db	168	GAATGACAGAACTAAGGAATATTTGGCTCTGATGAAGCCCTTCACTAAAGTCAATTTA	227
Qy	181	AAATTACCGACTGAAATGGCTGCTCTACATTCATTTATGTTGCTATTTGGGGCAAG	240
Db	228	AAATTACCGACTGAAATGGCTGCTCTACATTCATTTATGTTGCTATTTGGGGCAAG	287
Qy	241	AAATACATATTCGAACCTTTATGTTGAAGGGCTCCGCCATCTCGACTGACAGAAT	300
Db	288	AAATACATATTCGAACCTTTATGTTGAAGGGCTCCGCCATCTCGACTGACAGAAT	347
Qy	301	GGATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCGAATTTGATCACTTCTCTG	360
Db	348	GGATTTACAGCTTGCATTTAGCAGTTTACAGGATAATGCGAATTTGATCACTTCTCTG	407
Qy	361	CTTCACAGTGGAGCTGATATACAGCAGTTGGATACGGTGGCTCACTGCCCTCCATATT	420
Db	408	CTTCACAGTGGAGCTGATATACAGCAGTTGGATACGGTGGCTCACTGCCCTCCATATT	467
Qy	421	GCTACAAATAGCTGGCACTAGAGCTGCTGATGTGCTGTTGCAACATGGAGCTAATGTC	480
Db	468	GCTACAAATAGCTGGCACTAGAGCTGCTGATGTGCTGTTGCAACATGGAGCTAATGTC	527
Qy	481	AAATTTCAAGATCAGTTTCTTCTCACTCCATTCGATATTCGAGCTGCTATGGACATGAA	540
Db	528	AAATTTCAAGATCAGTTTCTTCTCACTCCATTCGATATTCGAGCTGCTATGGACATGAA	587
Qy	541	CAGGTAACTCGCCTCTTTTGAATTTGGTGTGCTGATGTAATGTAAGTGGTGAAGTTGGA	600
Db	588	CAGGTAACTCGCCTCTTTTGAATTTGGTGTGCTGATGTAATGTAAGTGGTGAAGTTGGA	647
Qy	601	GATAGACCCCTCACCTAGCATCTGCAAAAGATTTCTTGAATTTGCAAACTCTTGATG	660
Db	648	GATAGACCCCTCACCTAGCATCTGCAAAAGATTTCTTGAATTTGCAAACTCTTGATG	707
Qy	661	GAAGAGGCGCAAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCCTCCAT	720
Db	708	GAAGAGGCGCAAGCAAGCAGATGTAATGCTCAAGATAATGAAGACCATGTCCCCTCCAT	767
Qy	721	TTCTGTTCTCGATTTGGACACCATGATAGTTAAGTATCTGCTGCAAAAGTGAATTGGA	780
Db	768	TTCTGTTCTCGATTTGGACACCATGATAGTTAAGTATCTGCTGCAAAAGTGAATTGGA	827
Qy	781	GTTCAACCTCATGTTGTTAATCTATGAGATACCCCTTACCTGCTGATGCTACCAAT	840
Db	828	GTTCAACCTCATGTTGTTAATCTATGAGATACCCCTTACCTGCTGATGCTACCAAT	887
Qy	841	GGCAAAATTTGAAGTTGCAAGGAAATCATCCAAATATCAGGAACCAAGTCTGACTAAG	900
Db	888	GGCAAAATTTGAAGTTGCAAGGAAATCATCCAAATATCAGGAACCAAGTCTGACTAAG	947
Qy	901	GAAGACATCTTCAGTGAACAGCTTTTTCATAGTCTTGTACCTATGGCAGAGCATTCAC	960
Db	948	GAAGACATCTTCAGTGAACAGCTTTTTCATAGTCTTGTACCTATGGCAGAGCATTCAC	1007
Qy	961	CTAGTCAAAATTTCTTCTGATCAGATGTCATAAATCAACCAACCAAGGAGGATGGG	1020
Db	1008	CTAGTCAAAATTTCTTCTGATCAGATGTCATAAATCAACCAACCAAGGAGGATGGG	1067
Qy	1021	CACATCGAATTAACCTGCTGCTACCAAGCTGCTACATTCGGCTGGTTCAGTTCTTACTG	1080
Db	1068	CACATCGAATTAACCTGCTGCTACCAAGCTGCTACATTCGGCTGGTTCAGTTCTTACTG	1127

Qy	1081	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAATA	1140
Db	1128	GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCAGCAGGTCTAGTGGTGAATA	1187
Qy	1141	GATGACACACATGTTTGTATGTTGGCTTATGAAAAAGGGCATGATGCCATTTGTCACTC	1200
Db	1188	GATGACACACATGTTTGTATGTTGGCTTATGAAAAAGGGCATGATGCCATTTGTCACTC	1247
Qy	1201	CTGAAGCATTTAAGAGACCAAGATGAATTTGCCCTGTAAATGAATATCTCCAGCTGGA	1260
Db	1248	CTGAAGCATTTAAGAGACCAAGATGAATTTGCCCTGTAAATGAATATCTCCAGCTGGA	1307
Qy	1261	GGAGATGGCTCTATGTTGCTGTTCCATCCACCTTTGGGGAAGATTTAAAGCATGACAAA	1320
Db	1308	GGAGATGGCTCTATGTTGCTGTTCCATCCACCTTTGGGGAAGATTTAAAGCATGACAAA	1367
Qy	1321	GAGAAGCAGATATTTCTCCTCTAAGAGCTGGATTTGCCCTTCAATTTCCATCTTCAGTCT	1380
Db	1368	GAGAAGCAGATATTTCTCCTCTAAGAGCTGGATTTGCCCTTCAATTTCCATCTTCAGTCT	1427
Qy	1381	TCAGAAATTTGAGTTCCATGAGATTTTGGCTCAGGTCTCTTTGGGAAAAGTATATAAGGA	1440
Db	1428	TCAGAAATTTGAGTTCCATGAGATTTTGGCTCAGGTCTCTTTGGGAAAAGTATATAAGGA	1487
Qy	1441	CGATGAGAAATAAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCAAG	1500
Db	1488	CGATGAGAAATAAAATAGTGGCTATAAAACGTTATCGAGCCAAATACCTACTGCTCCAAG	1547
Qy	1501	TCAGATGTCGATATGTTTGGCGAGGCTGCCATTTCTGCCAGCTCAATCATCCCTGTC	1560
Db	1548	TCAGATGTCGATATGTTTGGCGAGGCTGCCATTTCTGCCAGCTCAATCATCCCTGTC	1607
Qy	1561	GTAATTCAGTTTGGTGGCTTGTGATGATCCAGCCAGTTTGGCAATTTGCTCACTCAA	1620
Db	1608	GTAATTCAGTTTGGTGGCTTGTGATGATCCAGCCAGTTTGGCAATTTGCTCACTCAA	1667
Qy	1621	TACATATCAGGGGGTCTCTGTTCTCCCTCTCTCATGAGCAGAGAGATTTCTTGATTTG	1680
Db	1668	TACATATCAGGGGGTCTCTGTTCTCCCTCTCTCATGAGCAGAGAGATTTCTTGATTTG	1727
Qy	1681	CAGTCTAAATTAATTTATTCAGTAGATTTGCAAGCATGGAGTACCTTCAACAACCTG	1740
Db	1728	CAGTCTAAATTAATTTATTCAGTAGATTTGCAAGCATGGAGTACCTTCAACAACCTG	1787
Qy	1741	ACACAGCCAAATTAACATCGTGACTTGAACAGTCAAAATTTCTCTCTATGAGGATGGG	1800
Db	1788	ACACAGCCAAATTAACATCGTGACTTGAACAGTCAAAATTTCTCTCTATGAGGATGGG	1847
Qy	1801	CATGCTGTGCTGCGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAAGACAAC	1860
Db	1848	CATGCTGTGCTGCGAGATTTTGGAGATCAAGATTTCTACAGTCTCTGATGAAGACAAC	1907
Qy	1861	ATGACAAAAACACCTGGGAAACCTCCGTTGGATGGCTCTCGAGGTGTTTCCAGCAGTCACT	1920
Db	1908	ATGACAAAAACACCTGGGAAACCTCCGTTGGATGGCTCTCGAGGTGTTTCCAGCAGTCACT	1967
Qy	1921	CGGTACACCATCAAGCAGATGTTTCACTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCT	1980
Db	1968	CGGTACACCATCAAGCAGATGTTTCACTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCT	2027
Qy	1981	GGCGAAATTTCCATTCGCTCATCTCAAGCCAGGGGCTGGCGAGCAGACATGGCTTACCAC	2040
Db	2028	GGCGAAATTTCCATTCGCTCATCTCAAGCCAGGGGCTGGCGAGCAGACATGGCTTACCAC	2087
Qy	2041	CACATCAGACCTCCCATTTGGCTATTTCCATTTCCAAAGCCCATATCATCTCTGCTGATAGA	2100
Db	2088	CACATCAGACCTCCCATTTGGCTATTTCCATTTCCAAAGCCCATATCATCTCTGCTGATAGA	2147
Qy	2101	GGGTGGAACGCTGCTCAAGGAGACCCGAAATTTTCTGAGTTTGTCTGATGAAGTTAGAA	2160
Db	2148	GGGTGGAACGCTGCTCAAGGAGACCCGAAATTTTCTGAGTTTGTCTGATGAAGTTAGAA	2207

QY 2161 GAGTGTCTCTCAACATTGAGCTGATGTCTCTGTCATCAAGTAAACAGAGTGGTCTCTC 2220
DB |||||
QY 2208 GAGTGTCTCTCAACATTGAGCTGATGTCTCTGTCATCAAGTAAACAGAGTGGTCTCTC 2267
DB |||||
QY 2221 TCACCTTCTCTCTCTCTGATTGCTGTGTAACCGGGAGAGCTGCGGAGTCAATG 2280
DB |||||
QY 2268 TCACCTTCTCTCTCTCTGATTGCTGTGTAACCGGGAGAGCTGCGGAGTCAATG 2327
DB |||||
QY 2281 GCAGCATTAAGAAGTCGTTTCGAATTGAATATATCTCTAAATGCAAGGTCCTATGCTGCT 2340
DB |||||
QY 2328 GCAGCATTAAGAAGTCGTTTCGAATTGAATATATCTCTAAATGCAAGGTCCTATGCTGCT 2387
DB |||||
QY 2341 TTGTCGCCAAAGTCGTGGAACAATATCTCTCAAGGTCGTCTTTGGAGGAGATGAAAAGA 2400
DB |||||
QY 2388 TTGTCGCCAAAGTCGTGGAACAATATCTCTCAAGGTCGTCTTTGGAGGAGATGAAAAGA 2447
DB |||||
QY 2401 AGTCTTCAATACACACCCATTGACAAATATGCTATGCTATCCGATCCCATGAGCTCAATG 2460
DB |||||
QY 2448 AGTCTTCAATACACACCCATTGACAAATATGCTATGCTATCCGATCCCATGAGCTCAATG 2507
DB |||||
QY 2461 CATTTTCATTCTCCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505
DB |||||
QY 2508 CATTTTCATTCTCCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2552
DB |||||

RESULT 5

AAS80720
ID AAS80720 standard; cDNA; 2788 BP.
AC AAS80720;
DT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #16524.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ASG16533.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 1; SEQ ID NO 16524; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (II) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2788 BP; 807 A; 585 C; 626 G; 770 T; 0 U; 0 Other;

Query Match 98.0%; Score 2455.4; DB 5; Length 2788;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2467; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 38 CTGATGAATCGAAGAAAAGTCAGTGAATCATATGTTATCAATAGAAAGATTAGAG 97
DB |||||
QY 317 CTGATGAATCGAAGAAAAGTCAGTGAATCATATGTTATCAATAGAAAGATTAGAG 376
DB |||||
QY 98 ATGACCTGCAGATCAAGGAAAAGAACTGCACAGAACTAAGGAATATATTGGCTCTGATG 157
DB |||||
QY 377 ATGACCTGCAGATCAAGGAAAAGAACTGCACAGAACTAAGGAATATATTGGCTCTGATG 436
DB |||||
QY 158 AAGCCTTCAGTAAAGTCAATTTAAATACCGCACTGAAAATGGGCTGTCTACTTCATT 217
DB |||||
QY 437 AAGCCTTCAGTAAAGTCAATTTAAATACCGCACTGAAAATGGGCTGTCTACTTCATT 496
DB |||||
QY 218 TATGTTCATTTGTGGAGGCAAGAAATCACATATTCGAACCTCTTATGTTGAAA-GGGGTC 276
DB |||||
QY 497 TATGTTCATTTGTGGAGGCAAGAAATCACATATTCGAACCTCTTATGTTGAAAAGGGGTC 556
DB |||||
QY 277 CGCCCATCTCGACTGCAAGAAATGGAATTCACAGCTTGCATTTACAGAGTTTACAAGAT 336
DB |||||
QY 557 CGCCCATCTCGACTGCAAGAAATGGAATTCACAGCTTGCATTTACAGAGTTTACAAGAT 616
DB |||||
QY 337 AATGCAGAAATGATCACTTCTCTGCTTCACAGTGGAGCTGATATACACAGCTTGATAC 396
DB |||||
QY 617 AATGCAGAAATGATCACTTCTCTGCTTCACAGTGGAGCTGATATACACAGCTTGATAC 676
DB |||||
QY 397 GGTGGCCTCACTGCCCTCCCATATTGCTACAATAGCTGGCCACCTAGAGGCTGCTGATG 456
DB |||||
QY 677 GGTGGCCTCACTGCCCTCCCATATTGCTACAATAGCTGGCCACCTAGAGGCTGCTGATG 736
DB |||||
QY 457 CTGTTGCAACATGGAGCTAATGTCATATTCAGATGCAATTTTCTTCTTCTTCTTCTTCTTCT 516
DB |||||
QY 737 CTGTTGCAACATGGAGCTAATGTCATATTCAGATGCAATTTTCTTCTTCTTCTTCTTCTTCT 796
DB |||||
QY 517 ATTGCAGCGTACTATGGACATGAACAGGTAACCTCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 576
DB |||||
QY 797 ATTGCAGCGTACTATGGACATGAACAGGTAACCTCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 856
DB |||||
QY 577 GTAAATGTAAGTGGTGAAGTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTC 636
DB |||||
QY 857 GTAAATGTAAGTGGTGAAGTGGAGATAGACCCCTCCACCTAGCATCTGCAAAAGGATTC 916
DB |||||
QY 637 TTGAATATTGCAAAACTCTTTGATGGAAGGAGGAGCAAGAGAGATGTGAATGCTCAAGAT 696
DB |||||
QY 917 TTGAATATTGCAAAACTCTTTGATGGAAGGAGGAGCAAGAGAGATGTGAATGCTCAAGAT 976
DB |||||
QY 697 AATGAAGACCATGTCCTCCACTCCATTTCTGTTCTCGAATTTGGACACCATGATATAGTTAAG 756
DB |||||
QY 977 AATGAAGACCATGTCCTCCACTCCATTTCTGTTCTCGAATTTGGACACCATGATATAGTTAAG 1036
DB |||||
QY 757 TATCTCTGCAAGTCAATTTGGAAGTTCACCTCATGTTGTTTAAATATCTATGGAGATACC 816
DB |||||
QY 1037 TATCTCTGCAAGTCAATTTGGAAGTTCACCTCATGTTGTTTAAATATCTATGGAGATACC 1096
DB |||||
QY 817 CCCTTACACTGGCATGCTACAAATGCAAAATTTGAAGTTGCCAAGGAAATCATCAATA 876
DB |||||
QY 1097 CCCTTACACTGGCATGCTACAAATGCAAAATTTGAAGTTGCCAAGGAAATCATCAATA 1156
DB |||||

QY 877 TCAGGAACAGAAAGTCTGACTAAGGAAACATCTTTCAGTGAACAGCTTTTCATAGTGCT 936
DB 1157 TCAGGAACAGAAAGTCTGACTAAGGAAACATCTTTCAGTGAACAGCTTTTCATAGTGCT 1216
QY 937 TGTACCTATGGCAAGAGCATGACCTAGTCAAAATTTCTTGTGATCAGAAATGTCATAAC 996
DB 1217 TGTACCTATGGCAAGAGCATGACCTAGTCAAAATTTCTTGTGATCAGAAATGTCATAAC 1276
QY 997 ATCAACACCAAGGAAGGATGGGACACCTGATTAACCTCTGCTTGTACACGGTCCAC 1086
DB 1277 ATCAACACCAAGGAAGGATGGGACACCTGATTAACCTCTGCTTGTACACGGTCCAC 1336
QY 1057 ATTTCGGCTTGTTCAGTCTTACTTGGATATGAGCTGATATGAATCTAGTGGCTTGTGAT 1116
DB 1337 ATTTCGGCTTGTTCAGTCTTACTTGGATATGAGCTGATATGAATCTAGTGGCTTGTGAT 1396
QY 1117 CCACAGAGTCTAGTGGTGAAGAGATGAGCAGACATGTTTGAATGAGCTTATGAAGAA 1176
DB 1397 CCACAGAGTCTAGTGGTGAAGAGATGAGCAGACATGTTTGAATGAGCTTATGAAGAA 1456
QY 1177 GGCAATGATGCCATTGTCACCTCTGAGCATTATAGAGACCAAGATGAATGGCC 1236
DB 1457 GGCAATGATGCCATTGTCACCTCTGAGCATTATAGAGACCAAGATGAATGGCC 1516
QY 1237 TGTAAATGAATATTTCTCAGCTGGAGAGATGGCTCTATGCTGTGTTCCATCACCCCTTG 1296
DB 1517 TGTAAATGAATATTTCTCAGCTGGAGAGATGGCTCTATGCTGTGTTCCATCACCCCTTG 1576
QY 1297 GGGAAGTAAAGAGATGACAAAGAGAGAGAGATATTTCTCTCTCTCTCTCTCTCTCTCT 1356
DB 1577 GGGAAGTAAAGAGATGACAAAGAGAGAGAGATATTTCTCTCTCTCTCTCTCTCTCTCT 1636
QY 1357 CCTTCACATTTCCATCTTCAGCTCTCAGAAATGAGTTCATGAGATTTATGGCTCAGGT 1416
DB 1637 CCTTCACATTTCCATCTTCAGCTCTCAGAAATGAGTTCATGAGATTTATGGCTCAGGT 1696
QY 1417 TCTTTTGGGAAGTATATAAGAGCAGATCGAGAAATAAATAGTGGCTATAAAGCTTAT 1476
DB 1697 TCTTTTGGGAAGTATATAAGAGCAGATCGAGAAATAAATAGTGGCTATAAAGCTTAT 1756
QY 1477 CGAGCAATACCTACTCTCAAGTCAGATGAGATGATGTTTTCGAGAGGTGCCAT 1536
DB 1757 CGAGCAATACCTACTCTCAAGTCAGATGAGATGATGTTTTCGAGAGGTGCCAT 1816
QY 1537 CTCTGCAGCTCAATCATCCCTCGTAAATTCAGTTTGGGTGCTTGTGTAATGATCCC 1596
DB 1817 CTCTGCAGCTCAATCATCCCTCGTAAATTCAGTTTGGGTGCTTGTGTAATGATCCC 1876
QY 1597 AGCCAGTTTGGCAATGCTCACTCAATACATATCAGGGGGTCTCTGTTCTCCCTCTTCAAT 1656
DB 1877 AGCCAGTTTGGCAATGCTCACTCAATACATATCAGGGGGTCTCTGTTCTCCCTCTTCAAT 1936
QY 1657 GACGAGAAGAGGATTTCTGATTTGCACTCTAAATTAATTTGCACTAGATGTTGCCAAA 1716
DB 1937 GACGAGAAGAGGATTTCTGATTTGCACTCTAAATTAATTTGCACTAGATGTTGCCAAA 1996
QY 1717 GGATGGAGTACCTTCACACCTGACACAGCAATTTATACATCGTGAATGAACAGTCCAC 1776
DB 1997 GGATGGAGTACCTTCACACCTGACACAGCAATTTATACATCGTGAATGAACAGTCCAC 2056
QY 1777 AATATTTCTCTCTATGAGGTGGCATGCTGTGGGAGATTTTGGAGATCAAGATTT 1836
DB 2057 AATATTTCTCTCTATGAGGTGGCATGCTGTGGGAGATTTTGGAGATCAAGATTT 2116
QY 1837 CTACAGTCTCTGGATGAAGACCAATGACAAACACCTCGGAACTCCGTTGGATGGCT 1896
DB 2117 CTACAGTCTCTGGATGAAGACCAATGACAAACACCTCGGAACTCCGTTGGATGGCT 2176
QY 1897 CCTGAGGTGTTCTCGAGTGCATCTCGGTACACCATCAAGAGAGATGTTCTAGCTATGCT 1956
DB 2177 CCTGAGGTGTTCTCGAGTGCATCTCGGTACACCATCAAGAGAGATGTTCTAGCTATGCT 2236

QY 1957 CTGTCTCTGTGGGAATTTCTCACTCGCGAAATTCATTCGCTCATCTCAAGCCAGCGCT 2016
DB 2237 CTGTCTCTGTGGGAATTTCTCACTCGCGAAATTCATTCGCTCATCTCAAGCCAGCGCT 2296
QY 2017 GGGCAGCAGACATGGCTTACCAACACATCAGACCTCCCATTTGGCTATTCATTTCCCAAG 2076
DB 2297 GGGCAGCAGACATGGCTTACCAACACATCAGACCTCCCATTTGGCTATTCATTTCCCAAG 2356
QY 2077 CCATATCATCTCTGCTGATAGAGGGTGGAAAGCATGCTCTGAAAGGAGACCGAATTT 2136
DB 2357 CCATATCATCTCTGCTGATAGAGGGTGGAAAGCATGCTCTGAAAGGAGACCGAATTT 2416
QY 2137 TCTGAAGTTGTCTCACTGAAGTTAGAAGAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGCA 2196
DB 2417 TCTGAAGTTGTCTCACTGAAGTTAGAAGAGTGTCTCTGCAACATTTGAGCTGATGCTCTCTGCA 2476
QY 2197 TCAAGTAAACAGCAGTGGGTCTCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2256
DB 2477 TCAAGTAAACAGCAGTGGGTCTCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2536
QY 2257 GGAGGACCTGGCGGAGTCACTGTGCGAGCATTTAAGAAGTTCGTTTCGAATTCGAATATGCT 2316
DB 2537 GGAGGACCTGGCGGAGTCACTGTGCGAGCATTTAAGAAGTTCGTTTCGAATTCGAATATGCT 2596
QY 2317 CTAAATGCAAGTCTCTATGCT 2376
DB 2597 CTAAATGCAAGTCTCTATGCT 2656
QY 2377 CTGTCTTTGGAGAGATGAAGAAGTCTTCAATATACACACCATTTGCAATATGCTCTCTCTCT 2436
DB 2657 CTGTCTTTGGAGAGATGAAGAAGTCTTCAATATACACACCATTTGCAATATGCTCTCTCTCT 2716
QY 2437 GTATCCGATCCCATGAGTCAATGCAATTTTCAATTTCTTCCGAAATAGTAGCAGCTTTGAG 2496
DB 2717 GTATCCGATCCCATGAGTCAATGCAATTTTCAATTTCTTCCGAAATAGTAGCAGCTTTGAG 2776
QY 2497 GACAGCAGC 2505
DB 2777 GACAGCAGC 2785

RESULT 6
AAA47609
ID AAA47609 standard; cDNA; 2505 BP.
XX AAA47609;
XX AC
XX XX
XX 20-OCT-2000 (first entry)
XX DE Rat CARK (Cardiac related Ankyrin-Repeat Protein Kinase) cDNA.
XX KW Cardiac related ankyrin repeat protein kinase; CARK; cytoskeleton;
XX KW cardiac cell growth factor receptor; Cell differentiation; modulator;
XX KW regulator; detection; cellular proliferation; cardiovascular disorder;
XX KW heart failure; hypertension; cancer; sarcoma; ds.
XX OS Rattus norvegicus.
XX XX
XX Key Location/Qualifiers
XX CDS 1..2505
XX FT /*tag= a
XX FT /product= "Rat CARK"
XX XX
XX WO200034330-A1.
XX PN
XX PD 15-JUN-2000.
XX XX
XX PF 10-DEC-1999; 99WO-US029465.
XX PR 11-DEC-1998; 98US-0111938P.
XX PR 14-APR-1999; 99US-00291839.
XX XX
XX (MILL-) MILLENNIUM PHARM INC.

XX Raju J;
PI WPI; 2000-431275/37.
DR New polynucleotide encoding cardiac-related ankyrin-repeat protein
XX kinase, useful for treating disorders such as cardiovascular disorders,
PT e.g. heart failure and cell differentiation disorders, e.g. cancer.
PT Claim 1; Page 155-159; 161pp; English.
XX
XX CARX polypeptides are regulators of signal transmission from cellular
CC receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton
CC function. They also act as modulators of cellular differentiation and
CC cell death. Nucleic acids, or their fragments encoding CARX polypeptides
CC are useful for detecting CARX nucleic acids especially mRNA, in a sample.
CC CARX is useful for treating disorders associated with upregulation or
CC downregulation of cellular proliferation such as, cardiovascular
CC disorders (heart failure and hypertension) and disorders associated with
CC cell differentiation such as cancer and sarcoma
XX
SQ Sequence 2505 BP; 648 A; 640 C; 617 G; 600 T; 0 U; 0 Other;
Query Match 75.5%; Score 1892.2; DB 3; Length 2505;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;
QY 1 ATGGGAATTATAATCTAGACCAACCCAACTTGTAAGTGAATGGAAGAAAGTC 60
DB 1 ATGGGAATTACAAATCCAGACCAACACAGACTTGTTCTGATCAATGGAAGAAAGTT 60
QY 61 AGTGATCATATGTTATACAAATAGAAAGATTAGAGATGACCTGCAGATCAAGGAAAA 120
DB 61 AGTGATCTTACGCTTATATCATAGAAAGGCTGGAGGATAACCTGCAGATCAAGGAAAT 120
QY 121 GAACCTGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCCTTCAGTAAAGTCAATTTA 180
DB 121 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGATGAAGCCTTCAGTAAAGTCAAGTTA 180
QY 181 AATTACCGCCTGAAATAGGCTGCTCTACTTCATTTATGTTGCAATTTGTGAGGCGAAG 240
DB 181 AATTACCGCAGAGCGTGGCTGCTCCCTGCTACACCTGTGCTGTCTGTGGCGGCAAC 240
QY 241 AAATCACATATTCGAATCTTATGTTGAAAGGGCTCGGCCATCTCGACTGACAGAAAT 300
DB 241 AAGTCATATTCGCTGCCCTTATGTTAAAGGGCTCGCTCCATCCAGACTGACGAGAAAT 300
QY 301 GGATTTACAGCCTTGCAATTTAGCAGTTTACAAGATAATGCAAGATTTGATCTCTCTG 360
DB 301 GGGTTTCCAGCTCTGCACCTGGCCGTTTACAAGGACAGCCCGGAACCTTATCACTTCACTG 360
QY 361 CTTACAGTGGAGTGATATACAGAGTTGGATACGGTGGCTCACTGGCTCCCATATT 420
DB 361 TTGCACAGCGAGCAGATGTTTCAGCAAGTGGGATACGGTGGCTTCACAGCCCTCCACATA 420
QY 421 GCTCAATAGCTGGACACCCAGAGGCTGCAGAAAGTGTGCTACAAATGGGCGCAATGTG 480
DB 421 AATATTCAAGATGACAGTTTTCATCCATTCGATATGACGGTACTATGCACATGAA 540
QY 481 AATGTTCAAGATGCGCTTCTTCACCCCACTGCACATTTGACGCCCTACTATGGGACGAG 540
QY 541 CAGGTAACAGTGCTCTTTTGAAGTTTGGTGTGCTGATGTAATTAAGTGGTGAAGTTGGA 600
DB 541 CAGGTAACAGTGCTCTTTTGAAGTTTGGTGTGCTGATGTAATTAAGTGGTGAAGTTGGA 600
QY 601 GATGAGCCCTTCACCTAGCATCTGCAAAAGGATCTTGAATTCGAAATCTCTGATG 660
DB 601 GACAGGCTCTGCACTGGCTCTGCAAAAGGCTCTTCAACATTTGAAATCTCTGATG 660
QY 661 GAAGAAGCAGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTGCCACTCCAT 720
DB 661 GAAGAAGCAGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTGCCACTCCAT 720

DB 661 GAAGAAGGAGCAAGAGCAGATGTGAACGCTCAGGACCAATGAAGAACAACGCTCCTCTGCAC 720
QY 721 TTCTGTTCTCGATTTGGACACCATGATATAGTAAATAGTATCTGCTGCAAAAGTGAATTTGAA 780
DB 721 TTCTGTTCTCGATTTGGACACCAACAATATAGTACGCTACCTGCTCAGAGTGAATTTAG 780
QY 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTACACTGCTGCACTGCTACAAAT 840
DB 781 GTCCAGCCTCAGCTCATTAACATCTATGCTGACCTCTTTTGCACCTGCTGCTACAAAT 840
QY 841 GGCATTTTCAAGTTGCCAAGGAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
DB 841 GGAATTTTGAAGTTGCCAAGGAATTTGCCAGTTAAACAGGAACCTGAAAGTCTGACTAAG 900
QY 901 GAAAAATCTTCAAGTGAAGACAGCTTTTCAATGCTGTTGTAATGCTGGAAGAGCAATTTGAC 960
DB 901 GAAAAATCTTCAAGTGAAGACAGCTTTTCAATGCTGTTGTAATGCTGGAAGAGCAATTTGAC 960
QY 961 CTAGTCAAAATTTCTTCTGATCAGATGTCATAAATCATCAACCAACCAAGGAAGGATGGG 1020
DB 961 CTGTCATAATTTCTTCTGATCAGAAATGCTGTGAACATTAACCAACGAGGAAGATGGG 1020
QY 1021 CACACTGGATTTACACTCTGCTTGTCTACCAAGCTTCACATTCGCTGCTGTTTCACTTCTG 1080
DB 1021 CACAGAGATTTGACCTCTGCTTGTCTACCAAGGCTATATCCGCTGCTGTTTCACTTCTG 1080
QY 1081 GATAATGGAGCTGATATGAATCTAGTGGCTTGTGATCCGAGCAGCTTAGTGGTGAAGAA 1140
DB 1081 GATAATGGTGCAGATATGAATCTTGTGCTTGTGATCCGAGCAGCTTAGTGGTGAAGAA 1140
QY 1141 GATGAGCAGACATGTTTGTGCTTATGAGGCTTATGAAAGGGCAGTATGCCATTTGTCCACATC 1200
DB 1141 GATGAGCAGACATGTTTGTGCTTATGAGGCTTACGAGGAAGGACATGATGCCATTTGTACATC 1200
QY 1201 CTGAGCATTTAAGAGACACCAAGATGAATTTGCTGCTGTAATGAATTTCTCAGGCTGGA 1260
DB 1201 CTGAGCATTTAAGAGACACCAAGATGAATTTGCTGCTGTAATGAATTTCTCAGGCTGGA 1260
QY 1261 GAGATGGCTCCTATGCTGCTGCTTCCATCACCTTGGGGAAGATTAAGAGATGACAAAA 1320
DB 1261 GAGATGGCTCCTATGCTGCTGCTTCCATCACCTTGGGGAAGATTAAGAGATGACAAAA 1320
QY 1321 GAGAAGCAGATATTTCT 1380
DB 1321 GAGAAGCAGATATTTCT 1380
QY 1381 TCAGAAATTTGAGTTCCATGAGATTTAGTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
DB 1381 TCAGAAATTTGAGTTCCAGGATTTAGTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
QY 1441 CGATGCAAGAAATAAAAATAGTGGCTATAAACGTTATTCAGGCAATACCTTCTGCTCCAA 1500
DB 1441 CGATGCAAGAAATAAAAATAGTGGCTATAAACGTTATTCAGGCAATACCTTCTGCTCCAA 1500
QY 1501 TCAGATGTTGATATGTTTTGCGAGAGGTTTCCATTTCTCTGCACTCAATCATCTCCCTGC 1560
DB 1501 TCAGATGTTGATATGTTTTGCGAGAGGTTTCCATTTCTCTGCACTCAATCATCTCCCTGC 1560
QY 1561 GTAAATTCAGTTTGTGGGCTGCTGCTGTAATGATCCAGCAGTTTGGCAATTTGCTCACTCAA 1620
DB 1561 GTGGTTTCAGTTTGTGGGCTGCTGCTGTAATGATCCAGCAGTTTGGCAATTTGCTCACTCAA 1620
QY 1621 TACATATCAGGGGGTCTCTGTTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
DB 1621 TACATATCAGGGGGTCTCTGTTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680
QY 1681 CAGTCTAAATTAATTAATTCAGTATGTTGCCAAGGATGAGTACCTTCAACACCTG 1740
DB 1681 CAGTCTAAATTAATTAATTCAGTATGTTGCCAAGGATGAGTACCTTCAACACCTG 1740
QY 1741 ACACAGCCAAATTAATCATCGTGAATTTGAACAGTCAATATTTCTTCTCTAGGATGGG 1800
DB 1741 ACCAGCCAAATTAATCATCGTGAATTTGAACAGTCAATATTTCTTCTCTAGGATGGG 1800

```

QY 1801 CATGCTGTGGTGGCAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
Db 1801 CATGCTGTGGTGGCAGATTTTGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAC 1860
QY 1861 ATACAAAACACCTGGGAACCTCGTTGGATGGCTCTCTGAGTGTTCACACAGTGCAC 1920
Db 1861 ATACAAAACACCTGGGAACCTCGTTGGATGGCTCTCTGAGTGTTCACACAGTGCAC 1920
QY 1921 CGGTACACCATCAAAAGCAGATGCTTTCAGTATGCTGTGTCTGTGGAAATTTCTCACT 1980
Db 1921 AGATACACCATCAAGGCTGATGCTTTCAGTATGCTGTGTCTGTGGAGCTCTCACT 1980
QY 1981 GGGGAAATTCATTCGCTCATCTCAAGCAGCGGCTCGGCGAGCAGATGCTGTACAC 2040
Db 1981 GGGGAAATTCATTCGCTCATCTCAAGCAGCGGCTCGGCGAGCAGATGCTGTACAC 2040
QY 2041 CACATCAGACCTCCCATTCGGTATTCATTCCTCAAGCCCATATCATCTCTGTATACGA 2100
Db 2041 CACATCAGACCTCCCATTCGGTATTCATTCCTCAAGCCCATCTCATCTCTGTATACGG 2100
QY 2101 GGGTGAACGCGATGCTCTGAAGAGACCCGAAATTTTCTGAAGTTGTCTAAGATTAGAA 2160
Db 2101 GGGTGAATGCGATGCTCTGAAGAGACCCGAGAGTTCTCTGAAGTCTGTAGCAAACTGGAG 2160
QY 2161 GAGTGTCTGCAACATTCGCTGATGCTCTCTCATCAAGTAACAGCAGTGGTCTCTC 2220
Db 2161 GAGTGTCTGCAACATTCGCTGATGCTCTCTCATCAAGTAACAGCAGTGGTCTCTC 2220
QY 2221 TCACCTTCTTCTTCTCTGATTCGCTGTGAACCGGGGAGACCTGGCGGAGTCATGTG 2280
Db 2221 TCACCTTCTTCTTCTCTGATTCGCTGTGAACCGGGGAGGCGCTGGCGGAGCACGTG 2280
QY 2281 GCAGCATTAAGAGTGTTCGAAATGGAAATGCTCTAAATGCAAGTCTCTATGCTGCT 2340
Db 2281 GCAGCCTTAAGAGTGTTCGAAATGGAAATGCTCTAAATGCAAGTCTCTATGCTGCT 2340
QY 2341 TTGTCCTCAAGTCTGCAATATTCCTCTCAAGTCTGTCTTTGGAGAGATGAAGA 2400
Db 2341 TTGTCCTCAAGTCTGCAATATTCCTCTCAAGTCTGTCTTTGGAGAGATGAAGA 2400
QY 2401 AGTCTTCAATACACCCATTCGAAATATGCTATGATTCGATCCCATGAGTCTAATG 2460
Db 2401 AGCACCAGTATTCAACTGTGCAAAATACGGCTATGCTGTATCCCATGAGCTGACG 2460
QY 2461 CATTTTCATCTTGGCGAAATAGTAGCAGCTTTGAGGACAGCAGC 2505
Db 2461 CACCTTCACTCCCGCAAGACGACGACCAACTTTGAGGACAGCAAC 2505

```

RESULT 7

AAA47608
ID AAA47608 standard; cDNA; 3026 BP.

XX AC AAA47608;
XX

DT 20-OCT-2000 (first entry)

DE Rat CARL (Cardiac related Ankyrin-Repeat Protein Kinase) cDNA.

XX Cardiac related ankyrin repeat protein kinase; CARL; cytoskeleton;
KW cardiac cell growth factor receptor; cell differentiation; modulator;
KW regulator; detection; cellular proliferation; cardiovascular disorder;
KW heart failure; hypertension; cancer; sarcoma; ds.

XX Rattus norvegicus.

XX PH Key Location/Qualifiers
FT CDS 61..2568

FT /*tag= a
FT /product= "Rat CARL"

XX WO200034330-A1.

XX 15-JUN-2000.
XX 10-DEC-1999; 99MO-US029465.
XX 11-DEC-1998; 98US-0111938P.
XX 14-APR-1999; 99US-00291839.
XX (MILL-) MILLENNIUM PHARM INC.
XX Raju J;
XX WPI; 2000-431275/37.
XX P-PSDB; AAB01474.
XX New polynucleotide encoding cardiac-related ankyrin-repeat protein
PT kinase, useful for treating disorders such as cardiovascular disorders,
PT e.g. heart failure and cell differentiation disorders, e.g. cancer.
XX
XX Claim 1; Fig 5; 161pp; English.
XX CARL polypeptides are regulators of signal transmission from cellular
CC receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton
CC function. They also act as modulators of cellular differentiation and
CC cell death. Nucleic acids, or their fragments encoding CARL polypeptides
CC are useful for detecting CARL nucleic acids especially mRNA, in a sample.
CC CARL is useful for treating disorders associated with upregulation or
CC downregulation of cellular proliferation such as, cardiovascular
CC disorders (heart failure and hypertension) and disorders associated with
CC cell differentiation such as cancer and sarcoma
XX
XX Sequence 3026 BP; 778 A; 751 C; 743 G; 754 T; 0 U; 0 Other;

Query Match 75.5%; Score 1892.2; DB 3; Length 3026;

Best Local Similarity 84.7%; Pred. No. 0;

Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

QY 1 ATGGGAAATTTATAATCTAGACCAACCCAACTTCTACTGATGATGGAAGAAAAGTC 60
Db 61 ATGGGAAATTTATAATCTAGACCAACCCAACTTCTACTGATGATGGAAGAAAAGTT 120
QY 61 AGTGAATCATATGTTATCATCAATGAAGAAGATTAGAAGTACCTGCAGATCAAGGAAAA 120
Db 121 AGTGAATCTTACGCTATTATCATGAAGAAGGCTGGAGGATAACCTGCAGATCAAGAAAAT 180
QY 121 GAATGACAGAACTAAGGAATATATTTGGCTCTGATGAAGCTTCAGTAAAGTCAATTTA 180
Db 181 GAATTTCAAGAACTAAGGCAATCTTTGGCTCTGATGAAGCTTCAGTGAAGTCAATTTA 240
QY 181 AATTACCGCACTGAATAATGGGCTGTCTTACTTCAATTTATGTTGCAATTTGGAAGGCAAG 240
Db 241 AATTACCGCACAGAGCGTGGCTGTCTCTGTCTACACCTCTGTGTCTGTGGCGGCAAC 300
QY 241 AATACATATTTGGAACCTTATGTTGAAGGGCTCCGCCATCTCGACTGCACAGAAAT 300
Db 301 AAGTCACATATCCGTGCCCTTAATGTTAAAGGGCTCCGCTCCATCCAGACTGACGAAAT 360
QY 301 GGATTTACAGCCTTGCAITTTAGCAGTTTACAAGGATAATGAGAAATGATCACTTCTCTG 360
Db 361 GGGTTTCCAGCTCTGCACCTGSCCGTTTACAAGGACAGCCCGGAACTTATCACTTCACTG 420
QY 361 CTTCAAGTGGAGCTGATATACAGCAGGTTGATACGGTGGCTCCTCACTGCCCTCATATT 420
Db 421 TTGCACAGCGGACAGATGTTTCAGCAAGTGGGATACGGTGGCTCCTCACAGCCTCCACATA 480
QY 421 GCTACAATAGTGGCCACCTCAGAGCTGCTGATGTGCTGTTGCAACATGGAGCTAAATGTC 480
Db 481 GCTGCAATAGCTGGACACCCAGAGCTGCAAGATGCTGCTACAACATGGGGCCCAATGTG 540
QY 481 AATATTCAAGATGCGAGTTTCTTTTCTACTCCATTGCAATATGAGCGGTACTATGGACATGAA 540
Db 541 AATGTTCAAGATGCGGCTCTTCTTCCACCCACTGCAATTTGACGCTTACTATGGGCAACGAG 600

QY 541 CAGGTAACCTCGCCTCTTTTGAATTTGGTGTGCTGATGTAATGTAAGTGGTGAAGTTGGA 600
DB |||||
DB 601 CAGGTAACCTCGCCTCTTTTGAATTTGGTGTGCTGATGTAATGTAAGTGGTGAAGTTGGA 660
QY 601 GATAGACCCCTCCACCTAGCATCTGCAAAAGGATTTTGAATATGCAAAACTCTTTGATG 660
DB |||||
DB 661 GACAGGCTCTGCACTGGCTCTGCAAGGGCTTTTCAACATTGTGAATCTCTGGTA 720
QY 661 GAAGAAGCAGCAAGCAGATGTGAATGTCTCAAGATTAATGAAGCAGATGCCATCTCAT 720
DB |||||
DB 721 GAAGAAGGAGCAAGCAGATGTGAACGCTCAGACCAATGAAGCAGCCTCCCTCTGCAC 780
QY 721 TCTGTCTCCATTTGGACACCATGATATAGTTAAGTATCTGCTGCAAAAGTATTTGGA 780
DB |||||
DB 781 TCTGTCTCCATTTGGACACCAATATATAGTGTGCTGCTCCAGATGACTTAGAG 840
QY 781 GTTCAACCTCATGTTGTTAATATCTATGGAGATACCCCTTTACACCTGGCATGCTACAAT 840
DB |||||
DB 841 GTCCAGCCTCACGTCATTAACTATCTATGGTGACACTCTTTTGACCTGGCATGCTACAAT 900
QY 841 GGCATAATTTGAAGTTGCCAGGAATCATCCAATATCAGGAACAGAAAGTCTGACTAG 900
DB |||||
DB 901 GGAATTTTGAAGTTGCCAGGAATTTGTCCAGGTAAACAGGAATCTGAAAAGTCTGACTAG 960
QY 901 GAAACATCTTCAGTGAACACAGCTTTTTCATAGTGTCTGTATCTATGGCAAGCATTGAC 960
DB |||||
DB 961 GAAACATCTTCAGGAGACAGCTTTTTCACAGTGTGTATCTATGGCAAGCATTGAC 1020
QY 961 CTAGTCAAAATTTCTTTGATCAGAAATGTCATAAATCAACATCAACCAAGAGAGGATGG 1020
DB |||||
DB 1021 CTGGTCAAAATTTCTTTGATCAGAAATGTCATAAATCAACATCAACCAAGAGAGATGG 1080
QY 1021 CACACTGATTAACACTCTGCTGCTACACAGGTCAACATTCGCTGCTGTTCACTTTACTG 1080
DB |||||
DB 1081 CACACAGATTTGCACTCTGCTGCTACACAGGCAATTCGCTGCTGTTCACTTTACTG 1140
QY 1081 GATAATGAGCTGATATGAATCTAGTGGCTGTGATCCAGAGAGTCTAGTGGTGAATAA 1140
DB |||||
DB 1141 GATAATGAGCTGATATGAATCTAGTGGCTGTGATCCAGAGAGTCTAGTGGTGAATAA 1200
QY 1141 GATGAGCAGACATGTTGATGCTGGCTTATGAAGAGGAGTATGCCATTTGTCACATC 1200
DB |||||
DB 1201 GATGAGCAGACATGTTGATGCTGGCTTACGAGAGGAGATGATGCCATTTGTCACATC 1260
QY 1201 CTGAAGCATTATAAGAGACCAAGATGAATGCTGCTGTAATGAATATTTCTCAGCCTGGA 1260
DB |||||
DB 1261 CTGAAGCATTATAAGAGACCAAGAGAGTGGAGTGGATGTAACGAATATTTCCAGCCTGGA 1320
QY 1261 GGAGATGGCTCTATGCTGTGTTCCATCACTTGGGAGAGATTAAGAGCATGACAAA 1320
DB |||||
DB 1321 GGAGATGGCTCTATGCTGTGTTCTTCCCTTGGGAGAGATTAAGAGCATGACAAA 1380
QY 1321 GAGAAGGAGATATTTCTCCTTAAGAGCTGGATTTGCCCTTCAATTTCCATCTTCAGCTC 1380
DB |||||
DB 1381 GAGAAGGAGATATTTCTCCTTCCCTCAGGCTGAACTACCTCCGCTTCCATCTCCACATC 1440
QY 1381 TCAGAAATTTGATTTCCATGAGATTAATTTGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
DB |||||
DB 1441 TCCGAAATTCGAGTTCCAGAGATTAATTCGGCTCGGGTTCTTTTGGGAAAGTCTATAAGGG 1500
QY 1441 CGATGCGAATAAATAAGTGGCTATAAAGAGTTTATCGAGCCAAATACCTACTGCTCCAAG 1500
DB |||||
DB 1501 CGATGCGAATAAATAAGTGGCTATAAAGAGTTTATCGAGCCAAATACCTACTGCTCCAAG 1560
QY 1501 TCAGATGAGATATTTTTCGAGAGGCTGCTCCATTTCTGCGAGCTCAATCATCCTGCG 1560
DB |||||
DB 1561 TCAGAGCTGGATATGTTTTCGAGAGGCTGCTCCATTTCTGCGAGCTCAACACCCCTGCG 1620
QY 1561 GTAATTCAGTTTGGGTGCTGCTGTAATGATCCAGCCAGTTTCCCATTTGCACTCAA 1620
DB |||||
DB 1621 GTGGTTCAGTTTGGGTGCTGCTGTAATGATCCAGCTGATGATCCAGCTGATTTGCTACTG 1680
QY 1621 TACATATCAGGGGGTCTCTGTTCTCTCCCTCTCTCATGAGCAGAGAGGATTTCTGATTG 1680

DB |||||
DB 1681 TACATTTCCAGAGGCTCCCTGTTCTCCCTGCTTTCATGAACAGAGAGATTTCTGACTTG 1740
QY 1681 CAGTCTAAATTAATTTAGTGTGATGTTGCCAAAGCATGGAGTACCTTCAACACTG 1740
DB |||||
DB 1741 CAGTCTAAATTAATTTAGTGTGATGTTGCCAAAGCATGGAGTACCTTCAACACTG 1800
QY 1741 ACACAGCAATTAATCATGTGATTTGAACAGTGTCAATATTTCTCTCTATGAGATGG 1800
DB |||||
DB 1801 ACCAGCAATCATACCCGACCTGGAACAGCAATATTTCTGCTCTATGAGATGGC 1860
QY 1801 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTACGTCTCTGATGAACAAAC 1860
DB |||||
DB 1861 CATGCTGTGTGGCAGATTTTGAGAAATCAAGATTTCTGCACTCTCTGATGAACAAAC 1920
QY 1861 ATGACAAACCACTGGGAACCTCGTGGATGGCTCTGAGGTGTTACGCGATGCACT 1920
DB |||||
DB 1921 ATGACAAAGCAGCAGGAACTCGCTGGATGGCTGGCTGGTGTTCACACAGTGCAG 1980
QY 1921 CGGTACACCATCAAGCAGATGTTTCAGTATGCTCTGCTGTCTGGGAAATTTCTCACT 1980
DB |||||
DB 1981 AGATACACCATCAAGCTGATGTTTCACTTACTCCCTGTGTCTGTGGAGCTCTCACT 2040
QY 1981 GGCAGAAATTCATTTCGCTCATCTCAAGCCAGCGGTGGGAGAGAGATGGCTTACCAAC 2040
DB |||||
DB 2041 GGAGAAATTCATTTCGCTCATCTCAAGCCAGCGGTGGGAGAGAGATGGCTTACCAAC 2100
QY 2041 CACATCAGACCTCCCATTTGGCTATTCATTCCTCAAGCCCATATCATCTCTGCTGATAGA 2100
DB |||||
DB 2101 CACATCAGACCGCCCATTCGGCTATTCATTCCTCCAGCCCATCTCATCTCTGCTGATAGA 2160
QY 2101 GGGTGAAGCAGCTGCTCTGAAAGAGACCCGAAATTTCTGAAGTTGTCTGAAGTTAGAA 2160
DB |||||
DB 2161 GGGTGAAGTGCATGCTCTGAAAGAGACCCAGAGTTCTCTGAAGTGTCTGACAACTGGAG 2220
QY 2161 GAGTGTCTGCAACATTTGAGCTGATGCTCTGCTCATCAAGTAAACAGCAGTGGGTCTCTC 2220
DB |||||
DB 2221 GAGTGTCTATGCAATTTGAGCTGATGCTCTCAGCATCAAGTAAACAGCAGTGGGTCTCTG 2280
QY 2221 TCACCTTCTTCTTCTGATTCCTGCTGAAACCCGAGGAGACCTGGCCGAGTCAATGTG 2280
DB |||||
DB 2281 TCACCTTCTTCTTCTGATTCCTGCTGAAACCCGAGGAGACCTGGCCGAGTCAATGTG 2340
QY 2281 GCAGCATTAAGAGTGTGTTTGAATTTGAAATGCTTAAATGCAAGTCTCTATGCTGCT 2340
DB |||||
DB 2341 GCAGCCTTACGAGCGCTTTTGAAGTTGGAGTATGCCCTTAAATGCAAGTCTCTATGCTGG 2400
QY 2341 TTGTCCCAAGTGTCTGACAAATATTCCTCTCAAGTGTCTGTTTGGAGAGATGAAAGA 2400
DB |||||
DB 2401 TGTTCCTCAAGTGTGGAACACACTTAATCCGGGCTGTCTTTGGAGAGATGAAATAGG 2460
QY 2401 AGTCTTCAATACACCCATTCGAAATATGCTATGATTCGATCCCATGAGTCAATG 2460
DB |||||
DB 2461 AGCAGCAGTATTAACACTGTTGCAAAATACGCTATGTTCTGATCCCATGAGCCTGAG 2520
QY 2461 CATTTTCATTTCTGCGGAATAGTAGCAGCTTTTGGAGCAGCAGC 2505
DB |||||
DB 2521 CACCTTCACTCCCGCAAGACGACGAACTTTGAGGACAGCAAC 2565

RESULT 8

ACC48581
ID ACC48581 standard; cDNA; 3026 BP.

XX ACC48581;

DT 11-AUG-2003 (first entry)

XX Rat cardiac-related ankyrin-repeat protein kinase (CARK) cDNA.

XX Rat; CARK; cardiac-related ankyrin-repeat protein kinase; enzyme;
KW cardiant; hypotensive; cytosolic; chromosome 1; gene therapy; gene; ss.

OS Rattus norvegicus.

XX Key Location/Qualifiers
 PH CDS 60..2568
 FT /*tag= a
 FT /product= "Rat CARP"
 FT /note= "the CDS is also claimed in Claim 1"
 XX

PN WO2003020912-A2.

XX

PD 13-MAR-2003.

XX 04-SEP-2002; 2002WO-US028300.

XX 05-SEP-2001; 2001US-00947199.

XX (MILL-) MILLENNIUM PHARM INC.

XX Raju J;

XX WPI; 2003-290188/28.

XX P-PSDB; ABR41887.

DR Novel isolated cardiac-related ankyrin-repeat protein kinase polypeptide,
 XX useful for treating cellular growth related disorders which include
 XX cardiovascular disorders and proliferative and/or differentiative
 XX disorders.

PS Claim 1; Fig 5A-D; 159pp; English.

CC The present sequence, the coding region of which is also claimed, is that
 CC of cDNA encoding a novel rat cardiac-related ankyrin-repeat protein
 CC kinase, designated CARP, which plays a role in signalling pathways
 CC associated with cardiac cell growth or differentiation. A plasmid
 CC containing the rat CARP nucleotide sequence is deposited as ATCC PTA-
 CC 1530. Human CARP gene (see ACC48580) expression is increased in ischaemic
 CC heart tissue samples, suggesting a role in the regulation of cardiac cell
 CC growth and/or differentiation and the pathogenesis of cardiovascular
 CC disorders, e.g. congestive heart failure and cardiac hypertrophy. The
 CC invention provides human and rat CARP nucleic acids, antisense molecules,
 CC recombinant expression vectors, host cells and transgenic animals in
 CC which a CARP gene has been introduced or disrupted, CARP proteins, fusion
 CC proteins, antigenic peptides and anti-CARP antibodies. These are useful
 CC as modulating agents for regulating a variety of cellular processes, e.g.
 CC cardiac cellular process, for modulating the phosphorylation state of a
 CC CARP molecule or one or more proteins involved in cellular growth or
 CC differentiation, for modulating cell behaviour or as targets and
 CC therapeutic agents controlling cardiac cell proliferation,

CC differentiation, hypertrophy and migration, for modulating intra- or inter
 CC cellular signalling and/or gene transcription, for modulating cell
 CC proliferation, growth, differentiation, survival and/or migration, for
 CC regulating transmission of signals from cellular receptors, for
 CC modulating entry of cells, e.g. cardiac precursor cells, into mitosis, or
 CC for regulating cytoskeletal function. The nucleic acids and proteins are
 CC useful for treating cellular growth related disorders which include
 CC cardiovascular disorders (such as heart failure, hypertension), and
 CC proliferative and/or differentiative disorders (such as cancer). They are
 CC also useful in screening assays, detection assays (e.g. chromosomal
 CC mapping, tissue typing, forensic biology), predictive medicine (e.g.,
 CC diagnostic assays, prognostic assays, monitoring clinical trials) and
 CC pharmacogenomics, and in methods of treatment. The nucleic acids are also
 CC useful as probes and primes, in gene therapy, in the detection of genetic
 CC alteration in the CARP gene, and in forensic identification

SQ Sequence 3026 BP; 778 A; 751 C; 743 G; 754 T; 0 U; 0 Other;

Query Match 75.5%; Score 1892.2; DB 7; Length 3026;
 Best Local Similarity 84.7%; Pred. No. 0;
 Matches 2122; Conservative 0; Mismatches 383; Indels 0; Gaps 0;

Qy 1 ATGGGAATTATAAATACACCAACCAACTTGCTACTGATGATGGAGAAAAGTC 60
 |||||
 Db 61 ATGGGGAATTACAAATCCAGACCAACACAGACTTGTCTGATGATGGAGAAAGT 120

Qy 61 AGTGAATCATATGTTTATCACAATGAAAGATTAGAAAGATGACCTGCAGATCAAGGAAAA 120
 Db 121 AGTGAATCTTACGCTATTATCATAGAAAGGCTGGAGGATACCTGCAGATCAAGAGAAAT 180
 Qy 121 GAATGACAGAACTAAGGAATATATTTGGCTCTGTATGAAGCTTCAGTAAAGTCAATTTA 180
 Db 181 GAATTTCAAGAACTAAGGCACATCTTTGGCTCTGTATGAAGCTTCAGTAAAGTCAATTTA 240
 Qy 181 AATTACCGCACTGAAATGCGGTCTCTACTTTCATTTATGTTGTCATTTGTTGGAGGCAAG 240
 Db 241 AATTACCGCAGAGAGGCTGGCTCTCCCTGTCTACACCTCTGCTGTGTCTGTGGGCGAAC 300
 Qy 241 AATACACATATTCGAACCTCTTATGTTGAAAGGCTCCGCCCATCTTCGATCGACAGAAAT 300
 Db 301 AAGTCACATATCCGTGCTCTTATGTTAAAGGGCTCCGTCCATCCAGACTGACGAGAAAT 360
 Qy 301 GGATTTACAGCCTTGCAATTTAGCAGTTTACAGGTAATGACAGATTCATCATTCTCTG 360
 Db 361 GGGTTTCCAGCTCTGCACCTGGCCGTTTACAAGGACAGCCCGGAACTTATCATTCTACTG 420
 Qy 361 CTTACAGTGGAGCTGATATACAGCAGGTTGGATACGGTGGCCTCACTGCCCTCCATATT 420
 Db 421 TTGCACAGCGGAGCAGATGTTTACGCAAGTGGGATACGGTGGCCTCACAGCCTCCACATA 480
 Qy 421 GCTACAATAGCTGGCCACCTAGAGGCTGCTGATGTCTGTGTTGCAACATGGAGCTAATGTC 480
 Db 481 GCTGCAATAGCTGGACACCCAGAGGCTGCAGAAGTGTCTGTACAAACATGGGGCCAAATGT 540
 Qy 481 AATATTCAAGATGTCAGTTTTTTTCTCACTCCATTCATATTGTCAGCGTACTATGGACATGAA 540
 Db 541 AATGTTCAAGATGCGCTCTTCTTCACTCCCATGTCACATTCGAGCCTACTATGGGCAAG 600
 Qy 541 CAGGTAACCTCGCCTCTTTTGAATTTGGTGTGTGATGTAATGTAAGTGGTGAAGTTGGA 600
 Db 601 CAGGTAACCAAGTGTCTTTTGAATTTGGTGTGTGATGTAATGTAAGTGGTGAAGTTGGA 660
 Qy 601 GATAGACCCCTCCACCTACATCTGCAAGGATTTCTTGAATTTGCAATTTGCAAACTCTTGATG 660
 Db 661 GACAGGCTCTGCACCTGGCCTCTTCAAGAGGCTCTTCAACATTTGTAATCTCTGTA 720
 Qy 661 GAAGAAGGAGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCACTCCAT 720
 Db 721 GAAGAAGGAGCAAGAGCAGATGTGAATGCTCAAGATAATGAAGACCATGTCCCTCTGCAC 780
 Qy 721 TTCTGTTCTCGATTGGACACCATGATATAGTAAATGATCTGCTCAAGATGATTGGA 780
 Db 781 TTCTGTTCTCGATTGGACACCATGATATAGTAAATGATCTGCTCAAGATGATTGGA 840
 Qy 781 GTTCAACCTCATGTTCTTAATATCTATGGAGATACCCCTTACACCTGGCATGCTTACAT 840
 Db 841 GTCCAGCCTCACGTCATTAACATCTATGTTGACATCTCTTTGCACCTGGCATGCTTACAT 900
 Qy 841 GGCAATTTTGAAGTTGCCAAGGAAATCATCCAAATATCAGGAACAGAAAGTCTGACTAAG 900
 Db 901 GGAAATTTTGAAGTTGCCAAGGAAATTTGCCAGGTAACAGGAACTGAAAGTCTGACTAAG 960
 Qy 901 GAAACATCTTCAGTGAACACAGCTTTTCATAGTGTCTGTACTATGGCAAGCATTCAC 960
 Db 961 GAAACATCTTCAGCAGACAGCTTTTTCATAGTGTCTGTACTATGGCAAGCATTCAC 1020
 Qy 961 CTAGTCAAAATTTCTTCTTGATCAGAAATGTCTATAACATCAACCCACCAAGGAGGATCGG 1020
 Db 1021 CTGCTCAAAATTTCTTCTTGATCAGAAATGTCTATAACATCAACCCACCAAGGAGATCGG 1080
 Qy 1021 CACACTGGATTCACCTCTGCTTGTACCAAGTGCATTCGCTGGTTCAGTCTTCTACTG 1080
 Db 1081 CACACAGGATTCACCTCTGCTTGTACCAAGGCTATATCCGCTGGTTCAGTCTTCTACTT 1140
 Qy 1081 GATAATGGAGCTCATATGAATCTAGTGGCTTGTGATCCAGCAGGCTCTAGTGGTGAAGAAA 1140
 Db 1141 GATATGGTGCAGATATGAATCTTGTGCTGTGATCCAGCAGGCTCTAGTGGTGAAGAAA 1200

PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
XX cancer or hepatitis.

XX Claim 5; Page 254; 282pp; English.

XX The invention relates to an isolated polypeptide, which is a human kinase
CC and phosphatase (KPP). KPP agonists and antagonists are useful for
CC diagnosing, treating or preventing disorders associated with aberrant
CC expression of KPP, particularly cell proliferative disorders (e.g.
CC arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal
CC nocturnal haemoglobinuria, polycythaemia vera, psoriasis, primary
CC thrombocytopaenia or cancer), developmental disorders (eg. renal tubular
CC acidosis, anaemia or mental retardation), neurological disorders (e.g.
CC Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/
CC inflammatory disorders (e.g. AIDS; acquired immune deficiency syndrome,
CC allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's
CC disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome,
CC gout, Grave's disease, Hashimoto's thyroiditis, irritable bowel syndrome,
CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
CC syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
CC bacterial, fungal, parasitic, protozoan or helminthic infections. The KPP
CC is useful in assessing the effects of exogenous compounds on the
CC expression of nucleic acids and kinases and phosphatases. KPP gene is
CC useful in gene therapy and for creating transgenic animals to model human
CC disease. The present sequence is human KPP cDNA

XX SQ Sequence 2024 BP; 613 A; 431 C; 425 G; 555 T; 0 U; 0 Other;

Query Match 70.7%; Score 1772; DB 8; Length 2024;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1775; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGGGAAATTATAAATCTAGACCAACCCAAACTGTACTGATGAATGGAAGAAAAAGTC 60
DB 46 ATGGGAAATTATAAATCTAGACCAACCCAAACTGTACTGATGAATGGAAGAAAAAGTC 105
QY 61 AGTGAATCATATGTTATCAATAGAAAAGATTAGAAAGATGACCTGCAGATCAAGAAAAA 120
DB 106 AGTGAATCATATGTTATCAATAGAAAAGATTAGAAAGATGACCTGCAGATCAAGAAAAA 165
QY 121 GAATGACAGAACTAAGGAATATTTGGCTGTGATGAAGCCCTTCAGTAAAGTCAATTTTA 180
DB 166 GAATGACAGAACTAAGGAATATTTGGCTGTGATGAAGCCCTTCAGTAAAGTCAATTTTA 225
QY 181 AATTACCGCACTGAAATGGCTGTCTCTACTTCAATTTATTTGTCATTTGTGGAGCAAG 240
DB 226 AATTACCGCACTGAAATGGCTGTCTCTACTTCAATTTATTTGTCATTTGTGGAGCAAG 285
QY 241 AATACATATTTGAACTCTTATGTTGAAGGGCTCCGCCATCTCGACTGACAGAAAT 300
DB 286 AATACATATTTGAACTCTTATGTTGAAGGGCTCCGCCATCTCGACTGACAGAAAT 345
QY 301 GGATTTACAGCCTTGCATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG 360
DB 346 GGATTTACAGCCTTGCATTTAGCAGTTTACAAGGATAATGCAAGATTGATCACTTCTCTG 405
QY 361 CTTACAGTGGAGCTGATATACAGAGGTTGATAGCGTGGCTCAGTCCCTCCATATT 420
DB 406 CTTACAGTGGAGCTGATATACAGAGGTTGATAGCGTGGCTCAGTCCCTCCATATT 465
QY 421 GCTACAACTAGTGGCCACCTAGAGGCTGTGATGTCGTGTCGATGCAATGGAGCTAATGTC 480
DB 466 GCTACAACTAGTGGCCACCTAGAGGCTGTGATGTCGTGTCGATGCAATGGAGCTAATGTC 525
QY 481 AATATTCAAGATCAGTCTTTTCTTCACTCCATGTCATATTGACGCGTACATGGAATGAA 540
DB 526 AATATTCAAGATCAGTCTTTTCTTCACTCCATGTCATATTGACGCGTACATGGAATGAA 585
QY 541 CAGTAACTCGCTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTCAAGTGA 600
DB 586 CAGTAACTCGCTCTTTTGAATTTGGTCTGATGTAATGTAAGTGGTCAAGTGA 645
QY 601 GATAGACCCCTCCACCTAGCATCTGCAAAAAGGATTCTTGAATATTGCAAAAACCTTGATG 660

DB 646 GATAGACCCCTCCACCTAGCATCTGCAAAAAGGATTCTTGAATATTGCAAAAACCTTGATG 705
QY 661 GAAGAAGGAGCAAGCAAGATGTAATGCTCAAGTAATGAAGACCATGTCCCACTCCAT 720
DB 706 GAAGAAGGAGCAAGCAAGATGTAATGCTCAAGTAATGAAGACCATGTCCCACTCCAT 765
QY 721 TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTCAAGTGAATTTGAA 780
DB 766 TTCTGTTCTCGATTTGGACACCATGATATAGTTAAGTATCTGCTCAAGTGAATTTGAA 825
QY 781 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTTACACCTGGCATGTACAA 840
DB 826 GTTCAACCTCATGTTGTTAATATCTATGAGATACCCCTTTACACCTGGCATGTACAA 885
QY 841 GGCAAATTTGAAGTTGCCAAGAAATCATCCAATATCAGGAACAGAAAGTCTGACTAAG 900
DB 886 GGCAAATTTGAAGTTGCCAAGAAATCATCCAATATCAGGAACAGAAAGTCTGACTAAG 945
QY 901 GAAAACATCTTCAGTGAAACAGCTTTTCATAGTGTGTTACCTATGCGCAAGAGCATTCAC 960
DB 946 GAAAACATCTTCAGTGAAACAGCTTTTCATAGTGTGTTACCTATGCGCAAGAGCATTCAC 1005
QY 961 CTAGTCAAAATTTCTTCTGATCAGAAATGTCAATAACATCAACCAAGAGGATGGG 1020
DB 1006 CTAGTCAAAATTTCTTCTGATCAGAAATGTCAATAACATCAACCAAGAGGATGGG 1065
QY 1021 CACACTGGATTACACTCTCTGTTGCTACACAGCTGCATTTGCGCTGGTTCAGTCTTACTG 1080
DB 1066 CACACTGGATTACACTCTCTGTTGCTACCAAGCTGCATTTGCGCTGGTTCAGTCTTACTG 1125
QY 1081 GATAATGGAGCTGATATGAATCTAGTGGCTGTGATGCCAGCAGGTCTAGTGGTGA AAAA 1140
DB 1126 GATAATGGAGCTGATATGAATCTAGTGGCTGTGATGCCAGCAGGTCTAGTGGTGA AAAA 1185
QY 1141 GATGACGACATGTTGATGTTGGCTTATGAAAAGGCGATGATGCCATGTGCACATC 1200
DB 1186 GATGACGACATGTTGATGTTGGCTTATGAAAAGGCGATGATGCCATGTGCACATC 1245
QY 1201 CTGAACATTTATAAGAGACCAAGATGAATTTGCGCTGTGAATGAAATTTCTCAGCCTGGA 1260
DB 1246 CTGAACATTTATAAGAGACCAAGATGAATTTGCGCTGTGAATGAAATTTCTCAGCCTGGA 1305
QY 1261 GGAGATGGCTCTATGTTGCTGTTCCATCACCCCTTTGGGGAAGATTAAAAGCATGCAAAA 1320
DB 1306 GGAGATGGCTCTATGTTGCTGTTCCATCACCCCTTTGGGGAAGATTAAAAGCATGCAAAA 1365
QY 1321 GAGAAGGCGAGATATTCTCTCTTAAGAGCTGGATTGCGCTTCAATTTCCATCTTCAGCTC 1380
DB 1366 GAGAAGGCGAGATATTCTCTCTTAAGAGCTGGATTGCGCTTCAATTTCCATCTTCAGCTC 1425
QY 1381 TCAGAAATTTGAGTTCCATGAGATTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1440
DB 1426 TCAGAAATTTGAGTTCCATGAGATTATGGCTCAGGTTCTTTTGGGAAAGTATATAAGGA 1485
QY 1441 CGATGAGAAATAAATAGTGGCTATAAACCTTATCGAGCCAACTACTGCTCCAAAG 1500
DB 1486 CGATGAGAAATAAATAGTGGCTATAAACCTTATCGAGCCAACTACTGCTCCAAAG 1545
QY 1501 TCAGATGTGGATATGTTTTCGCGAGAGGTGTCCATTTCTGCGAGCTCAATCATCCCTGC 1560
DB 1546 TCAGATGTGGATATGTTTTCGCGAGAGGTGTCCATTTCTGCGAGCTCAATCATCCCTGC 1605
QY 1561 GTAATTCAGTTTGTGGTGTCTGCTTGAATGATCCAGCCAGTTTGCCATGTGCACTCAA 1620
DB 1606 GTAATTCAGTTTGTGGTGTCTGCTTGAATGATCCAGCCAGTTTGCCATGTGCACTCAA 1665
QY 1621 TACATATCAGGGGGTTCCTCTGTTCTCCCTCTTCATGAGCAGAGAGGATTCTTGATTG 1680
DB 1666 TACATATCAGGGGGTTCCTCTGTTCTCCCTCTTCATGAGCAGAGAGGATTCTTGATTG 1725
QY 1681 CAGTCTAAATTAATTTGCAATGAGTGTGCAAGGATGAGATGCTTCACAACTG 1740

Db 1726 CAGTCTAAATTAATTATTGAGTAGATGTTGCCAAGGCGATGAGTACCTTCACAACCTG 1785
QY 1741 ACAGAGCAATATACATCGTGAATTCGAAAGAGTACACAATA 1780
Db 1786 ACAGAGCAATATACATCGTGAATTCGAAAGAGTACACAATA 1825

RESULT 10

ABT09088

ID ABT09088 standard; DNA; 668 BP.

XX AC

XX ABT09088;

XX DT 05-DEC-2002 (first entry)

XX DE Phase-1 Rat CT gene SEQ ID No 176.

XX KW Rat; toxicity study; rat toxic response gene; toxicological response;
XX KW drug development; phase-1 rat CT gene; ds.
XX OS Rattus sp.

XX PN WO200266682-A2.

XX PD 29-AUG-2002.

XX PF 29-JAN-2002; 2002WO-US002935.

XX PR 29-JAN-2001; 2001US-0264933P.

XX PR 26-JUL-2001; 2001US-0308161P.

XX PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.

XX PI Farris G, Hickman SH, Farr SB;

XX DR WPI; 2002-674961/72.

XX XX Evaluating the toxicity of an agent, useful in drug development or in
XX PT determining toxicological responses to a new drug, by determining the
XX PT expression of rat toxicologically relevant genes in the test animal in
XX PT response to the test agent.
XX PS Disclosure; Page 149; 388pp; English.

XX CC The invention relates to a method used for evaluating the toxicity of an
XX CC agent comprising determining the expression of a rat toxic response
XX CC gene(s) in the test animal in response to the agent. The method is useful
XX CC in drug development, particularly for conducting toxicity studies and
XX CC analysis before a new drug or compound is approved for human consumption
XX CC or use. The method is also useful in determining toxicological responses
XX CC to a new drug. This polynucleotide sequence represents a phase-1 rat CT
XX CC gene of the invention

SQ Sequence 668 BP; 162 A; 186 C; 170 G; 149 T; 0 U; 1 Other;

Query Match 18.6%; Score 466.4; DB 6; Length 668;
Best Local Similarity 87.1%; Pred. No. 7e-124;
Matches 512; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 1702 GTAGATGTTGCCAAGGCGATGGAGTACCTGCACAGCTTGACCAAGCAATTTATACATGCT 1761

Db 77 GTAGACGTTGCCAAGGCGATGGAGTACCTGCACAGCTTGACCAAGCAATTTATACACCGC 136

QY 1762 GACTTGACAGTCAATATTTCTCTATGAGATGGGCGATGCTGTGGTGCAGATTTT 1821

Db 137 GACTTGACAGTCAATATTTCTCTATGAGATGGGCGATGCTGTGGTGCAGATTTT 196

QY 1822 GGAGATCAAGATTTCTACAGTCTCTCGATGAAGACAAATGCAAAACAACTGGGAAC 1881

Db 197 GGAGATCAAGATTTCTGACGCTCTCGATGAAGACAAATGCAAAACAACTGGGAAC 256

QY 1882 CTCGTTGGATGGCTCTGAGTGTTCACGAGTGCATCGGTACACCATCAAGCAGAT 1941

Db 257 CTGCGTGGATGGCCCTGAGGTGTTTCACACAGTGCACGAGATACACCATCAAGCTGAT 316
QY 1942 GTCTTCAGCTATGCTGTGTCTGTGGGAAATTCACCTGGGAAATTCATTTGCTCAT 2001
Db 317 GTCTTCAGTACTCCCTGTGTGTGGGAGCTCCTCCTGGAGAAATTCATTTGCTCAT 376
QY 2002 CTCAAGCCAGCGGCTGGGCGAGCAGACATGGCTTACCAACCATCAGACCTCCCATGGC 2061
Db 377 CTCAAGCCAGCGGCTGGGCGAGCAGACATGGGCTATCACCACATCAGACCGCCATGGC 436
QY 2062 TATTCATTCGCAAGCCCATATCATCTCTGTGTATACGAGGTGGAAGCATGCTCTGAA 2121
Db 437 TATTCATTCGCAAGCCCATATCATCTCTGTGTATACGAGGTGGAAGCATGCTCTGAA 496
QY 2122 GGAAGACCCGAATTTCTGAAGTTGTCATGAAGTTAGAAAGTGTCTCTGCAACATTCAG 2181
Db 497 GGAGCAGCAGATTTCTCTGAAGTCGTTAGCAAACTGGAGAGTGCCTATGCAATGTGAG 556
QY 2182 CTGATGTCCTGTCATCAAGTAAACAGCAGTGGGTCTCTCTACCTTCTTCTTCTGAT 2241
Db 557 CTCATGTCCTCAGCATCAAGTAAACAGCAGTGGCTCTCTGTACCTTCTTCTTCTCCGAT 616
QY 2242 TGCTGTGTGAACCGGCGAGGACCTGCGCGAGTCTATGTGGCAGCATTA 2289
Db 617 TGCTGTGTGAACCGGCGAGGACCTGCGCGAGGACCTGCGCGAGGACCTTAA 664

RESULT 11

AAS33234

ID AAS33234 standard; cDNA; 633 BP.

XX AC AAS33234;

XX DT 04-DEC-2001 (first entry)

XX DE DNA encoding human secreted protein, Seq ID No 193.

XX KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
XX KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
XX KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
XX KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
XX KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
XX KW Gaucher's disease; neurological disease; cerebrovascular disorder;
XX KW thrombosis; wound healing; ss.

XX OS Homo sapiens.

XX PN WO200155326-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001347.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-02321968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 23-SEP-2000; 2000US-0234997P.
PR 23-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244517P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 03-JAN-2001; 2001US-0259678P.
XX
XX
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-451931/48.
P-PSDB; AAU20525.
New nucleic acids and polypeptides, useful for diagnosing, preventing or treating medical conditions.
Claim 1; SEQ ID NO 193; 753pp; English.
The invention relates to novel isolated nucleic acid molecules (I) encoding human secreted proteins (II). (I) and (II) are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in the prevention, treatment and diagnosis of diseases associated with inappropriate expression of secreted proteins. (I) and complementary sequences may also be used as DNA probes in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect and quantitate the presence of similar nucleic acid sequences in samples, and so which patients may be in need of restorative therapy. (II) may also be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of the expression and activity of the secreted proteins. The anti-(II) antibodies and antagonists may also be used to down regulate expression and activity of (II). The anti-(II) antibodies may also be used as diagnostic agents for detecting the presence of (II) in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The disorders include for example: immune/autoimmune diseases (e.g. HIV

PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244817P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-451931/48.
DR P-PSDB; AAU20648.
XX

PT New nucleic acids and polypeptides, useful for diagnosing, preventing or
treating medical conditions.

PS Claim 1; SEQ ID NO 316; 753pp; English.

XX The invention relates to novel isolated nucleic acid molecules (I)
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
CC the prevention, treatment and diagnosis of diseases associated with
CC inappropriate expression of secreted proteins. (I) and complementary
CC sequences may also be used as DNA probes in diagnostic assays (e.g.
CC polymerase chain reactions (PCR)) to detect and quantitate the presence
CC of similar nucleic acid sequences in samples, and so which patients may
CC be in need of restorative therapy. (II) may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of the secreted
CC proteins. The anti-(II) antibodies and antagonists may also be used to
CC down regulate expression and activity of (II). The anti-(II) antibodies
CC may also be used as diagnostic agents for detecting the presence of (II)
CC in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The
CC disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular
CC disorders (e.g. corneal infections). (I) and (II) agonists, antagonists
CC and antibodies can also be used to promote wound healing, maintain organs
CC and before transplantation, and support cell culture of primary tissues.

Query Match 17.2%; Score 432; DB 4; Length 640;

Best Local Similarity 97.9%; Pred. No. 66-114;

Matches 457; Conservative 1; Mismatches 7; Indels 2; Gaps 2;

Qy	909	CTTCAGTGAACAGCTTTTCATAGTGTGCTTGTACCTATGCGAAGACGATGACCTAGTCAA	968
Db	44	CTTCATGAACACAGCTTTTCATAGTGTGCTTGTACCTATGCGAAGACGATGACCTAGTCAA	103
Qy	969	ATTTCTTCTTGATCAGATGTCTATAAATCAACCAAGAGGATGGGCACACTGG	1028
Db	104	ATTTCTTCTTGATCAGATGTCTATAAATCAACCAAGAGGATGGGCACACTGG	163
Qy	1029	ATTACACTCTGCTTGCTACACGGTCAATTCCTGCTGTTTCTTACTGGATAATGG	1088
Db	164	ATTACACTCTGCTTGCTACACGGTCAATTCCTGCTGTTTCTTACTGGATAATGG	223
Qy	1089	AGCTGATATGATCTAGTGGCTTGATCCCGAGAGCTAGTGGTGAAGAAGATGACA	1148
Db	224	AGCTGATATGATCTAGTGGCTTGATCCCGAGAGCTAGTGGTGAAGAAGATGACA	283
Qy	1149	GACATCTTTTGATGTGGGCTTATCAAAAAGGGCATGATCCATTGTACACTCTCTGAAGCA	1208
Db	284	GACATCTTTTGATGTGGGCTTATCAAAAAGGGCATGATCCATTGTACACTCTCTGAAGCA	343
Qy	1209	TTATAAGACACCAAGATGAATTCCTGCTGAATTAATTTCTAGCCTGGAGAGATGG	1268
Db	344	TTATAAGACACCAAGATGAATTCCTGCTGAATTAATTA-TCTCAGCTGGAGAGATGG	402
Qy	1269	CTCCTATGTCTGTTTCCATCACCTTGGGGAAGATTAAAGCATGACAAAGAGAAGGC	1328
Db	403	CTCCTATGTCTGTTTCCATCACCTTGGGGAAGATTAAAGCATGACAAAGAGAAGGC	462
Qy	1329	AGATATCTCCTCCTAA-GAGCTGGATGCTTCCATTTCCATCTT	1374

```

DB      463 AGATATTCCTCTCTAGGAGCTGGATTGCTTCACATTCATCT 509
|||||
RESULT 13
AAS80722
ID      AAS80722 standard; cDNA; 1557 BP.
XX
AC      AAS80722;
XX
DT      13-FEB-2002 (first entry)
XX
DE      DNA encoding novel human diagnostic protein #16526.
XX
KW      Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW      food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS      Homo sapiens.
XX
PN      WO200175067-A2.
XX
PD      11-OCT-2001.
XX
PF      30-MAR-2001; 2001WO-US008631.
XX
PR      31-MAR-2000; 2000US-00540217.
PR      23-AUG-2000; 2000US-00649167.
XX
PA      (HYSE-) HYSEQ INC.
XX
PI      Dmanac RT, Liu C, Tang YT;
XX
XX      WPI; 2001-639362/73.
DR      P-PSDB; ABG16535.
XX
XX      New isolated polynucleotide and encoded polypeptides, useful in
PT      diagnostics, forensics, gene mapping, identification of mutations
PT      responsible for genetic disorders or other traits and to assess
PT      biodiversity.
XX
PS      Claim 1; SEQ ID NO 16526; 103pp; English.
XX
CC      The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC      sequences. (I) is useful as hybridisation probes, polymerase chain
CC      reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC      and in recombinant production of (II). The polynucleotides are also used
CC      in diagnostics as expressed sequence tags for identifying expressed
CC      genes. (I) is useful in gene therapy techniques to restore normal
CC      activity of (II) or to treat disease states involving (II). (II) is
CC      useful for generating antibodies against it, detecting or quantitating a
CC      polypeptide in tissue, as molecular weight markers and as a food
CC      supplement. (II) and its binding partners are useful in medical imaging
CC      of sites expressing (II). (I) and (II) are useful for treating disorders
CC      involving aberrant protein expression or biological activity. The
CC      polypeptide and polynucleotide sequences have applications in
CC      diagnostics, forensics, gene mapping, identification of mutations
CC      and to produce other types of data and products dependent on DNA and
CC      amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC      coding sequences of the invention. Note: The sequence data for this
CC      patent did not appear in the printed specification, but was obtained in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences
XX
SQ      Sequence 1557 BP; 417 A; 400 C; 459 G; 281 T; 0 U; 0 Other;
Query Match      5.4%; Score 135.6; DB 5; Length 1557;
Best Local Similarity 68.4%; Pred. No. 4.8e-28;
Matches 245; Conservative 0; Mismatches 59; Indels 54; Gaps 2;
XX
QY      40 GATGAATCGAAGAAAAAGTCAAGTCAATCATATGTTATCAATAGAAAGATTAGAGAT 99
DB      178 GATGAATCGAAGAAAAAGTCAAGTCAATCATATGTTATCAATAGAAAGATTAGAGAT 237
|||||
100 GACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAGGAATATATTTGGCTCTGATGAA 159
|||||
238 GACCTGCAGATCAAGGAAAAAGAACTGACAGAACTAAGGAATATATTTGGTGGTTTGTG 297
|||||
160 GCCTT-----CAGTAA 171
|||||
298 ATCTTGTGCTTTCAGGAGTGAAGCTGCAGGCTTCATGCTGAGTGTATACAGCTCATAAA 357
|||||
172 GTCAATTTAAATTACCGCACTGAAATGGCTGTCTCTACTTCA-----TTTATGTTTC 225
|||||
358 GGCAGTCCGACCCCAAGAGTGAAGAGCAGCAAGATTATTGCAAAAAGCGAAGACAA 417
|||||
226 ATTTGTGGAGGCAAGAAATACATATTCGAACTCTTATGTTGAAAGGGCTCGGCCATCT 285
|||||
418 AGCTCCACCGCAAGAAATCACATATTCGAACTCTTATGTTGAAAGGGCTCGGCCATCT 477
|||||
286 CGACTGACAAAGAAATGGAATTTACAGCTTGCATTTAGCAGTTTACAAGGATAATGCAG 343
|||||
478 CGACTGACAAAGAAATGGAATTTACAGCTTGCATTTAGCAGTTTACAAGCCGAGGAG 535
|||||
RESULT 14
AAC48526
ID      AAC48526 standard; DNA; 2283 BP.
XX
AC      AAC48526;
XX
DT      18-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana DNA fragment SEQ ID NO: 57797.
XX
KW      Hybridisation assay; genetic mapping; gene expression control;
KW      protein identification; signal transduction pathway; metabolic pathway;
KW      promoter; termination sequence; ss.
XX
OS      Arabidopsis thaliana.
XX
PN      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-00301439.
XX
PR      25-FEB-1999; 99US-0121825P.
PR      05-MAR-1999; 99US-0123180P.
PR      09-MAR-1999; 99US-0123548P.
PR      23-MAR-1999; 99US-0125788P.
PR      25-MAR-1999; 99US-0126264P.
PR      29-MAR-1999; 99US-0126785P.
PR      01-APR-1999; 99US-0127462P.
PR      06-APR-1999; 99US-0128234P.
PR      08-APR-1999; 99US-0128714P.
PR      16-APR-1999; 99US-0129845P.
PR      19-APR-1999; 99US-0130077P.
PR      21-APR-1999; 99US-0130449P.
PR      23-APR-1999; 99US-0130510P.
PR      28-APR-1999; 99US-0130891P.
PR      30-APR-1999; 99US-0131449P.
PR      30-APR-1999; 99US-0132048P.
PR      04-MAY-1999; 99US-0132407P.
PR      05-MAY-1999; 99US-0132484P.
PR      06-MAY-1999; 99US-0132485P.
PR      07-MAY-1999; 99US-0132486P.
PR      07-MAY-1999; 99US-0132487P.
PR      11-MAY-1999; 99US-0132863P.
PR      14-MAY-1999; 99US-0134218P.
PR      14-MAY-1999; 99US-0134219P.
PR      14-MAY-1999; 99US-0134221P.
PR      14-MAY-1999; 99US-0134370P.
PR      18-MAY-1999; 99US-0134768P.
PR      19-MAY-1999; 99US-0134941P.

```

PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151103P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158322P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159333P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.

PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 4.2%; Score 106.4; DB 3; Length 2283;
Best Local Similarity 52.3%; Pred. No. 1.6e-19; Indels 18; Gaps 4;
Matches 339; Conservative 0; Mismatches 291;

QY 1516 TTTTCCGAGAGGTGTCATCTCTGCCAGCTCAATCATCCCTGCGTAATTCAGTTTGTG 1575
DB TTTTCTCAAGAAGTTTTTATAATGAGGAAGTTCGACACAAAACGTGTTCAATTTTG 1041
QY 1576 GGTGCTTCTGTAATGATCCAGCAGTTTGCATGTCTCACTCAATACATACAGGGGT 1635
DB 1042 GGTGATGCACAAGATCTCCAACCTCTGT---ATAGTGACTGAGTTATGGCTCGAGG 1098
QY 1636 TCTGTGTTCTCCCTCTTCATGACGAGAAGGATCTTGATTTGCAGTCTAAATTAAT 1695
DB 1099 AGCATATATGATTTCTTCAACACAGAAATGCGTTTCAAACTTCAACTTACTCAA 1158
QY 1696 ATTGCAGTATGTTGCAAGGATGAGTACCTTCAACCTGACAGCCAAATATA 1755
DB 1159 GTTGCACTGTATGTCGAAAAGGAATGAGCTATTGTCATCA-----AAACAATTA 1212
QY 1756 CATGCTGACTTGAACGTCAATATTTCTCTATGAGGATGGCATGCTGTGGTGCA 1815
DB 1213 CACAGGACCTTAGACTGGAATCTTCTATGATGGAACATGAGTTGTCAGGTTGCT 1272
QY 1816 GATTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGCAACATGACAAACCT 1875
DB 1273 GATTTGCGAGTTGCCAGAGT-----GCAGATTGAATCAGGGTCTATGCTGAACT 1326
QY 1876 GGAACCTCCGTTGGATGGCTCTGAGGTGTTCAAGCAGTGCATCGGTACACCATCAA 1935
DB 1327 GGGACATACCGGTGGATGGCTCCAGAGG---TCATTGAGCAAAACCTTACATCAAG 1383
QY 1936 GCAGATGCTTTCAGCTATGCTCTGTCTGTGGAAATTTCTCACTGCGGAAATTCATTC 1995
DB 1384 GCAGATGCTTTCAGTTATGATATGCTATGGAATCTCTGACTGGTGACATCCCATAT 1443
QY 1996 GCTCATCTCAAGCAGCGGTGCGGACGACAGATGGCTTACACACATCAGACTCCC 2055
DB 1444 GCTTCTTCTGACTCAACAGCAGCTGTTGGCGTTGTCCAAAAGGGGCTTCGACCCAAA 1503
QY 2056 ATTGGCTATTCCATTCCCAAGCCCATATCTCTGCTGATACGAGGTGGAACGATGT 2115
DB 1504 ATCCCAAGAAACACACCCAAAAGTGAAGGGCTTCTAGAGAGATGCTGGCATCAAG 1563
QY 2116 CTGAAGGAGACCGGAATTTTCTGAAGTTGTCATGAAGTTAGAAG 2163
DB 1564 CCAGAACAGACCACTGTTTGAGGAATCATAGAAATGCTACAACAG 1611

RESULT 15
ID ADA69514 standard; DNA; 1704 BP.
XX ADA69514;
AC ADA69514;
DT 20-NOV-2003 (first entry)
DE Rice gene, SEQ ID 2837.
XX plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX Oryza sativa.
XX WO2003000998-A1.
XX 03-JAN-2003.
PD

XX 22-JUN-2001; 2001WO-IB001105.
PR 22-JUN-2001; 2001WO-IB001105.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX Claim 6; SEQ ID NO 2837; 899pp; English.
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX Sequence 1704 BP; 463 A; 344 C; 479 G; 417 T; 0 U; 1 Other;
Query Match 4.0%; Score 100.2; DB 7; Length 1704;
Best Local Similarity 52.1%; Pred. No. 8.6e-18;
Matches 328; Conservative 0; Mismatches 283; Indels 18; Gaps 4;
QY 1516 TTTTCCGAGAGGTGTCATCTCTGCCAGCTCAATCATCCCTGCGTAATTCAGTTTGTG 1575
DB 982 TTTGCCGAGAGATATATATATGAGGAAGTTCGCCACAGAAATGTTGTACATTTATT 1041
QY 1576 GGTGCTTGTGTAATGATCCAGCCAGTTTGCCATTGTCTCACTCAATATCATCGGGGT 1635
DB 1042 GGTGATGC---ACTAAACCCCAATCTATGTATGTACAGAAATATATGTCAGTGGG 1098
QY 1636 TCTGTGTTCTCCCTCTTCATGAGCAGAGAGATTCCTGATTTGCAGTCTTAAATTAAT 1695
DB 1099 AGTGTGTTATGATTTACCTGCAATAAACAATAAAGGTGTTATCAAGCTCCCTGTTTCTTGA 1158
QY 1696 ATTGCAGTATGTTGCCAAAGGCATGGAGTACCTTCAACCTGACACAGCCAAATATA 1755
DB 1159 GTTGTGATGATGTTTCAAAAGGCATGACTACTTGCACCA-----AAACAATTAAT 1212
QY 1756 CATGCTGACTTGAACAGTCAATATTTCTTCTATGAGGATGGGATGCTGTGGTGCA 1815
DB 1213 CATCGATTTGAAAACCTGCTATCTCTCTTATGGATGAAAATGGGACAGTTAAGTTGCT 1272
QY 1816 GATTTGGAGAAATCAAGATTTCTACAGTCTCTGGATGAAGACAAATGACAAAACAACT 1875
DB 1273 GATTTGGTGTTCAGCTGTTTAAAGCTCAATCTGGAGTA-----ATGACTGCTGAACA 1326
QY 1876 GGAACCTCCGTTGGATGGCTCCTGAGGTGTTTACGAGTGCATCGGTACACCATCAA 1935
DB 1327 GGGACATATCGTTGGATGGCTCCAGAGGTTATAGAACA---CAAGCCCTATGATCACAAG 1383
QY 1936 GCAGATGCTTTCAGCTATGCTCTGTCTGTGGAAATTTCTCACTGCGGAAATTCATTC 1995
DB 1384 GCGATGTTTAAAGTTTGAATTTTGAATTTGATGTTGGAGCTCTCAAGGAGATTCCTTAT 1443
QY 1996 GCTCATCTCAAGCAGCGGTGCGGACGACAGATGGCTTACCAACATCAGACTCCC 2055
DB 1444 GAGTACCTGACACCATTTACAAGCAGCTGTAGGTGTGGTTTCAAGAGGATTACGGCTACA 1503
QY 2056 ATTGGCTATTCCATTTCCAGGCCATATCATCTCTGCTGATACGAGGTGGAACGATGT 2115

Db 1504 ATACCAAAATGCTCATGCAAACTTCTGAGCTTCTTCAGAAATGTTGGCAACAGGAG 1563
Qy 2116 CCTGAGGAAAGACCCGAATTTCTGAAGT 2144
Db 1564 CCTGCTGAAAGACCAGATTTCTCTGAAT 1592

Search completed: September 6, 2004, 01:05:54
Job time : 636.459 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 2, 2004, 14:07:41 ; Search time 136 Seconds
(without alignments)
1734.759 Million cell updates/sec

Title: US-10-626-173-2

Perfect score: 4390
Sequence: 1 MGNYKSRPTQCTDEWKKV.....PMSSMHFHSRNSSSFFDSS 835

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Genesep_29Jan04:*

- 1: Genesep1980s:*
- 2: Genesep1980s:*
- 3: Genesep2000s:*
- 4: Genesep2001s:*
- 5: Genesep2002s:*
- 6: Genesep2003as:*
- 7: Genesep2003bs:*
- 8: Genesep2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4390	100.0	835	3 AAB01470	Aab01470 Human CAR
2	4390	100.0	835	4 AAB65674	Aab65674 Novel pro
3	4390	100.0	835	6 ABR41886	Abr41886 Human car
4	4311	98.2	928	4 ABG16533	Abg16533 Novel hum
5	4079	92.9	835	3 AAB01474	Aab01474 Rat CAR
6	4079	92.9	835	6 ABR41887	Abr41887 Rat cardi
7	3102	70.7	603	6 AAE37962	Aae37962 Human kin
8	603	13.7	148	4 AAU20525	Aau20525 Human sec
9	603	13.7	148	4 AAU20648	Aau20648 Human sec
10	493	11.2	589	3 AAG45984	Aag45984 Arabidops
11	493	11.2	732	3 AAG45983	Aag45983 Arabidops
12	493	11.2	760	3 AAG45982	Aag45982 Arabidops
13	460.5	10.5	1881	7 ADD47763	Add47763 Human pro
14	454	10.3	1536	6 ABU11523	Abu11523 Human MDD
15	446	10.2	1549	4 ABB67412	Abb67412 Drosophil
16	446	10.2	1549	4 ABB58328	Abb58328 Drosophil
17	442.5	10.1	1056	6 AAE33684	Aae33684 Human str
18	440.5	10.0	1053	4 AAU28174	Aau28174 Novel hum
19	440	10.0	705	4 AAU20496	Aau20496 Human sec
20	438	10.0	412	3 AAG22170	Aag22170 Arabidops
21	437.5	10.0	369	3 AAG22172	Aag22172 Arabidops
22	437.5	10.0	374	3 AAG22171	Aag22171 Arabidops
23	435	9.9	489	4 ABUS3032	Abus3032 Cell stru
24	435	9.9	1839	7 ADD27862	Add27862 Human bra
25	435	9.9	1872	4 AAM79160	Aam79160 Human pro

ALIGNMENTS

RESULT 1

AAB01470 AAB01470 standard; protein; 835 AA.

XX AAB01470;

XX AC

XX 20-OCT-2000 (first entry)

XX Human CAR (Cardiac related Ankyrin-Repeat Protein Kinase).

XX Cardiac related ankyrin repeat protein kinase; CAR; cytoskeleton;

XX cardiac cell growth factor receptor; cell differentiation; modulator;

XX regulator; detection; cellular proliferation; cardiovascular disorder;

XX heart failure; hypertension; cancer; sarcoma.

XX Homo sapiens.

XX WO200034330-A1.

XX 15-JUN-2000.

XX 10-DEC-1999; 99WO-US029465.

XX 11-DEC-1998; 98US-0111938P.

XX 14-APR-1999; 99US-00291839.

XX (MILL-) MILLENNIUM PHARM INC.

XX Raju J;

XX WPI; 2000-431275/37.

XX N-PSDB; AAA47606, AAA47607.

XX New polynucleotide encoding cardiac-related ankyrin-repeat protein kinase, useful for treating disorders such as cardiovascular disorders, e.g. heart failure and cell differentiation disorders, e.g. cancer.

XX Claim 2; Fig 1; 161pp; English.

XX CARP polypeptides are regulators of signal transduction from cellular receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton function. They also act as modulators of cellular differentiation and cell death. Nucleic acids, or their fragments encoding CARP polypeptides are useful for detecting CARP nucleic acids especially mRNA, in a sample. CARP is useful for treating disorders associated with upregulation or downregulation of cellular proliferation such as, cardiovascular disorders (heart failure and hypertension) and disorders associated with cell differentiation such as cancer and sarcoma

26 433 9.9 407 3 AAG32051 Arabidops
27 432 9.8 425 5 AAM52832 Physcomit
28 431.5 9.8 447 4 ABUS3031 Cell stru
29 428.5 9.8 2622 7 ADE55508 Rat Prote
30 427.5 9.7 367 5 AAG32053 Arabidops
31 427.5 9.7 367 5 ABB83477 Human cyt
32 427.5 9.7 369 3 AAG32052 Arabidops
33 427.5 9.7 1088 4 AAB62429 Human kid
34 427.5 9.7 1094 6 AAE33688 Human str
35 427.5 9.7 1097 4 AAG21941 Novel hum
36 427.5 9.7 4274 4 AAG00972 Novel hum
37 427.5 9.7 4386 4 ABG07375 Novel hum
38 426 9.7 448 4 ABB65101 Drosophil
39 422.5 9.6 1762 7 ADB85236 Rat 190 k
40 419 9.5 518 7 ABR56190 Human 163
41 419 9.5 567 2 AAY28998 Human TGF
42 419 9.5 579 2 AAW27092 Mouse tra
43 419 9.5 579 2 AAY28996 Human TGF
44 419 9.5 579 5 ABB85033 Pain regu
45 419 9.5 506 2 AAY28997 Human TGF


```
Db 241 FCSRFGRHDIKYLQSLDLEVPQVNVNIGDTPHLACYNKGKFEVAKELIQISGESLTK 300
Qy 301 ENIFSETAFHSACTYCKSIDLKFLLDQNVININHGDRGHTGLHSACVHGHIRLVOELL 360
Db 301 ENIFSETAFHSACTYCKSIDLKFLLDQNVININHGDRGHTGLHSACVHGHIRLVOELL 360
Qy 361 DNGADNMLVACDPSRSGGKDEQTCMLWYKXGHDIAIVTLLKHYPQDELPCNEYSQPG 420
Db 361 DNGADNMLVACDPSRSGGKDEQTCMLWYKXGHDIAIVTLLKHYPQDELPCNEYSQPG 420
Qy 421 GDGSYVSPPLGKIKSMYKEXKADILLRAGLPSHPLQLSIEFHEIIGSGFQKVKYG 480
Db 421 GDGSYVSPPLGKIKSMYKEXKADILLRAGLPSHPLQLSIEFHEIIGSGFQKVKYG 480
Qy 481 RCNKKVAIKRYRANTYCKSDVDMFCREVSILCOLNHPCVIOFVGACLNPSQFAIVTQ 540
Db 481 RCNKKVAIKRYRANTYCKSDVDMFCREVSILCOLNHPCVIOFVGACLNPSQFAIVTQ 540
Qy 541 YISGSLFSLHQRKILDLQSKLIITAVDAKGMELHNLTOPIIHRDLNHNILLYEDG 600
Db 541 YISGSLFSLHQRKILDLQSKLIITAVDAKGMELHNLTOPIIHRDLNHNILLYEDG 600
Qy 601 HAVVADFGESRFLQSDENMTKQPCNLRWMAPEVFTQCTRTIKADVFSVALCLWEILT 660
Db 601 HAVVADFGESRFLQSDENMTKQPCNLRWMAPEVFTQCTRTIKADVFSVALCLWEILT 660
Qy 661 GEIPFAHLKFAAAADWYHHRPPIGYSIPKPISSLIRGNWACPEGRPFSEVVMKLE 720
Db 661 GEIPFAHLKFAAAADWYHHRPPIGYSIPKPISSLIRGNWACPEGRPFSEVVMKLE 720
Qy 721 ECLCNTELSPASSNSGSLSSSSSDCLVNRGGGRSHVAALRRFRLEVALNARSYAA 780
Db 721 ECLCNTELSPASSNSGSLSSSSSDCLVNRGGGRSHVAALRRFRLEVALNARSYAA 780
Qy 781 LSQAGYSSQGLSLEBMKRSIQYTPIDKYGVVSPMSMHFHSRNSSSPEDSS 835
Db 781 LSQAGYSSQGLSLEBMKRSIQYTPIDKYGVVSPMSMHFHSRNSSSPEDSS 835

RESULT 3
ABR41886
ID ABR41886 standard; protein; 835 AA.
XX
XX
AC ABR41886;
XX
XX
DT 11-AUG-2003 (first entry)
DE
XX
XX
XX Human cardiac-related ankyrin-repeat protein kinase (CARK).
KW Human; CARK; cardiac-related ankyrin-repeat protein kinase; enzyme;
KW Cardiant; hypertensive; cytostatic.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 2..7
FT Modified-site /note= "N-myristoylation motif"
FT Modified-site 11..14
FT Modified-site /note= "potential casein kinase II phosphorylation site"
FT Modified-site 18..21
FT /note= "potential cAMP/cGMP dependent protein kinase phosphorylation site"
FT Modified-site 50..55
FT /note= "N-myristoylation motif"
FT Domain 66..99
FT /note= "Ankyrin repeat domain"
FT Modified-site 78..83
FT Modified-site /note= "N-myristoylation motif"
FT Modified-site 91..96
FT /note= "N-myristoylation motif"
FT Domain 100..132
FT /note= "Ankyrin repeat domain"
FT Modified-site 123..126

/note= "potential casein kinase II phosphorylation site"
133..165
/note= "Ankyrin repeat domain"
133..138
/note= "N-myristoylation motif"
157..162
/note= "N-myristoylation motif"
168..198
/note= "Ankyrin repeat domain"
130..195
/note= "N-myristoylation motif"
194..197
/note= "potential N-glycosylation site"
199..233
/note= "Ankyrin repeat domain"
208..210
/note= "potential protein kinase C phosphorylation site"
224..227
/note= "potential casein kinase II phosphorylation site"
234..268
/note= "Ankyrin repeat domain"
257..260
/note= "potential casein kinase II phosphorylation site"
259..302
/note= "Ankyrin repeat domain"
293..296
/note= "potential casein kinase II phosphorylation site"
306..338
/note= "Ankyrin repeat domain"
339..371
/note= "Ankyrin repeat domain"
343..348
/note= "N-myristoylation motif"
363..368
/note= "N-myristoylation motif"
376..379
/note= "potential casein kinase II phosphorylation site"
411..415
/note= "LXCXE motif"
420..425
/note= "N-myristoylation motif"
461..464
/note= "potential casein kinase II phosphorylation site"
463..716
/note= "protein kinase domain"
469..477
/note= "protein kinase ATP binding region"
499..502
/note= "potential casein kinase II phosphorylation site"
526..531
/note= "N-myristoylation motif"
545..550
/note= "N-myristoylation motif"
579..582
/note= "potential N-glycosylation site"
615..618
/note= "potential casein kinase II phosphorylation site"
620..623
/note= "potential N-glycosylation site"
643..645
/note= "potential protein kinase C phosphorylation site"
735..738
/note= "potential N-glycosylation site"
744..747
/note= "potential casein kinase II phosphorylation site"
764..771
/note= "potential tyrosine kinase phosphorylation site"
765..768
/note= "potential casein kinase II phosphorylation site"
794..797
/note= "potential casein kinase II phosphorylation site"
805..808
/note= "potential casein kinase II phosphorylation site"
```

FT Modified-site 824..826 /note="potential protein kinase C phosphorylation site"
 FT Modified-site 827..830 /note="potential N-glycosylation site"
 FT Modified-site 829..832 /note="potential casein kinase II phosphorylation site"
 FT Modified-site 876..791 /note="N-myristoylation motif"
 FT WO2003020912-A2.
 PN 13-MAR-2003.
 XX 04-SEP-2002; 2002WO-US028300.
 XX 05-SEP-2001; 2001US-00947199.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Raju J;
 XX WPI; 2003-290188/28.
 DR N-PSDB; ACC48580.
 XX Novel isolated cardiac-related ankyrin-repeat protein kinase polypeptide,
 PT useful for treating cellular growth related disorders which include
 PT cardiovascular disorders and proliferative and/or differentiative
 PT disorders.
 XX Claim 14; Fig 1B; 158pp; English.
 PS The present sequence is the protein sequence of a novel 96 kDa human
 CC cardiac-related ankyrin-repeat protein kinase, designated CARX, which
 CC plays a role in signalling pathways associated with cardiac cell growth
 CC or differentiation. CARX is a dual specificity protein kinase having
 CC specificity for both serine/threonine and tyrosine residues. CARX gene
 CC expression is increased in ischaemic heart tissue samples, suggesting a
 CC role in the regulation of cardiac cell growth and/or differentiation and
 CC the pathogenesis of cardiovascular disorders, e.g. congestive heart
 CC failure and cardiac hypertrophy. The invention provides CARX proteins,
 CC fusion proteins, antigenic peptides and anti-CARX antibodies, and also
 CC CARX nucleic acids, antisense molecules, recombinant expression vectors,
 CC host cells and transgenic animals. These are useful as modulating agents
 CC for regulating a variety of cellular processes, e.g. cardiac cellular
 CC process, for modulating the phosphorylation state of a CARX molecule or
 CC one or more proteins involved in cellular growth or differentiation, for
 CC modulating cell behaviour or as targets and therapeutic agents
 CC controlling cardiac cell proliferation, differentiation, hypertrophy and
 CC migration for modulating intra- or inter-cellular signalling and/or gene
 CC transcription, for modulating cell proliferation, growth,
 CC differentiation, survival and/or migration, for regulating transmission
 CC of signals from cellular receptors, for modulating entry of cells, e.g.
 CC cardiac precursor cells, into mitosis, or for regulating cytoskeletal
 CC function. The proteins and nucleic acids are useful for treating cellular
 CC growth related disorders which include cardiovascular disorders (such as
 CC heart failure, hypertension), and proliferative and/or differentiative
 CC disorders (such as cancer). They are also useful in screening assays,
 CC detection assays (e.g. chromosomal mapping, tissue typing, forensic
 CC biology), predictive medicine (e.g., diagnostic assays, prognostic
 CC assays, monitoring clinical trials) and pharmacogenomics, and in methods
 CC of treatment. A CARX polypeptide can be used to generate a specific
 CC antibody, to screen for naturally-occurring CARX substrates, and to
 CC screen for drugs or compounds which modulate CARX activity. It is also
 CC useful as a bait protein in a yeast two-hybrid or three-hybrid assay and
 CC for identifying other proteins which bind to, or interact with, CARX or
 CC which are involved in CARX activity
 XX Sequence 835 AA;
 Query Match 100.0%; Score 4390; DB 6; Length 835;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGNYKSRPTCTCTDEWKKKVSSEYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL 60
 DB 1 MGNYKSRPTCTCTDEWKKKVSSEYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL 60
 QY 61 NYRTENGLSLHLHCCICGGKKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDNABELITSL 120
 DB 61 NYRTENGLSLHLHCCICGGKKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDNABELITSL 120
 QY 121 LHSAGDIQOVGYGGLTALHTIATAGLEADVLLQHGANNVIOQAVFFTPPLHIAAYGYHE 180
 DB 121 LHSAGDIQOVGYGGLTALHTIATAGLEADVLLQHGANNVIOQAVFFTPPLHIAAYGYHE 180
 QY 181 QVTRLLKFKGADVNVSGEVDRLPLHLASAKGFNIAKLMEEGSKADVNAQDNEDHVPFLH 240
 DB 181 QVTRLLKFKGADVNVSGEVDRLPLHLASAKGFNIAKLMEEGSKADVNAQDNEDHVPFLH 240
 QY 241 FCSRFHHDIKYLLOSLDLEVOHPVNIYGDTPHLACYNKGFVAKELIIQISGTESITK 300
 DB 241 FCSRFHHDIKYLLOSLDLEVOHPVNIYGDTPHLACYNKGFVAKELIIQISGTESITK 300
 QY 301 ENTFSETAFHSACTYKGSIDLKFLDQNVININHOGRDGHGHTGLHSACVHGHIRLVQFLL 360
 DB 301 ENTFSETAFHSACTYKGSIDLKFLDQNVININHOGRDGHGHTGLHSACVHGHIRLVQFLL 360
 QY 361 DNGADNMLVACDPSRSGEKDEQTCMLMAYEKGHDAIVTLKHYYKRPQDELPCNYSQPG 420
 DB 361 DNGADNMLVACDPSRSGEKDEQTCMLMAYEKGHDAIVTLKHYYKRPQDELPCNYSQPG 420
 QY 421 GDGSYVSPSPGLKIKSMTEKADIILLRAGLPSHFHLQLSIEIFHEIIGSGSFQKVKYG 480
 DB 421 GDGSYVSPSPGLKIKSMTEKADIILLRAGLPSHFHLQLSIEIFHEIIGSGSFQKVKYG 480
 QY 481 RCRNKIVAIKRYRANTYCSKSDVDMFCREVSVILCOLNHPFCVIOFVGACLNDSQFAIVTQ 540
 DB 481 RCRNKIVAIKRYRANTYCSKSDVDMFCREVSVILCOLNHPFCVIOFVGACLNDSQFAIVTQ 540
 QY 541 YISGGSLSLHLHCKRILDLQSKLIITAVDVAKGMEYLNLTQPIIHRDLNSHNILLYEDG 600
 DB 541 YISGGSLSLHLHCKRILDLQSKLIITAVDVAKGMEYLNLTQPIIHRDLNSHNILLYEDG 600
 QY 601 HAVVADFGESRFLQSLDEDNMTKQGNLWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660
 DB 601 HAVVADFGESRFLQSLDEDNMTKQGNLWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660
 QY 661 GEIPFAHLKPAALAAADMAVHHIRPPIGYIPKPISSILLIRGNVACPEGEPESEVVMKLE 720
 DB 661 GEIPFAHLKPAALAAADMAVHHIRPPIGYIPKPISSILLIRGNVACPEGEPESEVVMKLE 720
 QY 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPQSHVAALRSRFELEYALNARSYAA 780
 DB 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPQSHVAALRSRFELEYALNARSYAA 780
 QY 781 LSQAGAGYSSQGLSLEEMKRSLOYTIDKYGVYVDPMSMHFHSRCNSSFDDSS 835
 DB 781 LSQAGAGYSSQGLSLEEMKRSLOYTIDKYGVYVDPMSMHFHSRCNSSFDDSS 835
 RESULT 4
 ABG16533
 ID ABG16533 standard; protein; 928 AA.
 XX ABG16533;
 AC ABG16533;
 XX 18-FEB-2002 (first entry)
 DT Novel human diagnostic protein #16524.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 XX WO200175067-A2.
 PN

XX 11-OCT-2001.
ZD
XX 30-MAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR
XX 23-AUG-2000; 2000US-00649167.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
XX N-PSDB; AAS80720.
DR
XX New isolated polynucleotide and encoded polypeptides, useful in
FT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
FT biodiversity.
XX
PS Claim 20; SEQ ID NO 46892; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 928 AA;
Query Match 98.2%; Score 4311; DB 4; Length 928;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 821; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
13 TDWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGDEAFSKVNLNRYTENGSLH 72
106 SDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGDEAFSKVNLNRYTENGSLH 165
73 LCCICGKXSHITMLKGLRPSRLTRNGFTALHLAVYKDNALITSLHSGADIQQVGY 132
166 LCCICGKXSHITMLKGLRPSRLTRNGFTALHLAVYKDNALITSLHSGADIQQVGY 225
133 GGLTALHIAIAGHLEAADVLLQHGANNVIOQAVFPTPLHIAAYYGHEQVTRLLKFGAD 192
226 GGLTALHIAIAGHLEAADVLLQHGANNVIOQAVFPTPLHIAAYYGHEQVTRLLKFGAD 285
193 VNVSGEVDRPLHLASAKGFLNTAKLMBEGSKADVNAQDNEDHVPFLHFCSPFGHDIIVK 252
286 VNVSGEVDRPLHLASAKGFLNTAKLMBEGSKADVNAQDNEDHVPFLHFCSPFGHDIIVK 345
253 YLLQSDLEVPVHVNIYGDTPHLACVNGKFEVAXEIIQISGTESLTKENIFSETAFHSA 312
346 YLLQSDLEVPVHVNIYGDTPHLACVNGKFEVAXEIIQISGTESLTKENIFSETAFHSA 405
313 CTYGKSIDLVKFLDDQNVININHQGRDGHGTGHSACYHGHIKRLVQFLDNGADNMLVACD 372
406 CTYGKSIDLVKFLDDQNVININHQGRDGHGTGHSACYHGHIKRLVQFLDNGADNMLVACD 465

QY 373 PRRSGEKDEQTCMLWAYEKGHDAIVTLLKHKKRPODELPCEYKSOQGGDSYVSPPL 432
DB 466 PRRSGEKDEQTCMLWAYEKGHDAIVTLLKHKKRPODELPCEYKSOQGGDSYVSPPL 525
QY 433 GIKSMTEKADILLRAGLPSHFHLQSEIEFHEIIIGSGSFGKVKYKGRCKNKIVAIKRY 492
DB 526 GIKSMTEKADILLRAGLPSHFHLQSEIEFHEIIIGSGSFGKVKYKGRCKNKIVAIKRY 585
QY 493 RANTYCSKSDVDMFCREVSILCOLNHPCVIQVGACLNDPQOFALVTOYISGGSLSFLH 552
DB 586 RANTYCSKSDVDMFCREVSILCOLNHPCVIQVGACLNDPQOFALVTOYISGGSLSFLH 645
QY 553 EOKRIIDLOSLLIIADVAKGMEYLHNLTOTPIHRDLNSHNILLYEDGHVAVVDFGESRF 612
DB 646 EOKRIIDLOSLLIIADVAKGMEYLHNLTOTPIHRDLNSHNILLYEDGHVAVVDFGESRF 705
QY 613 LQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFYALCLMEILLTGEIPFAHLKPA 672
DB 706 LQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFYALCLMEILLTGEIPFAHLKPA 765
QY 673 AADNAYHHRPIGYSIPKIPISLLIRGNACPRGPRFESEVVMKLEECNIELMSPA 732
DB 766 AADNAYHHRPIGYSIPKIPISLLIRGNACPRGPRFESEVVMKLEECNIELMSPA 825
QY 733 SSNSGSLSPSSSDCLVNRGGPRSHVAALRSRFELEYALNARSYAALSQAGYSSQG 792
DB 826 SSNSGSLSPSSSDCLVNRGGPRSHVAALRSRFELEYALNARSYAALSQAGYSSQG 885
QY 793 LSLEEMKRSLOQTPIDKYGVSDPMSSMHFHSRCNSSFEDSS 835
DB 886 LSLEEMKRSLOQTPIDKYGVSDPMSSMHFHSRCNSSFEDSS 928
RESULT 5
AAB01474
ID AAB01474 standard; protein; 835 AA.
XX AC AAB01474;
XX DT 20-OCT-2000 (first entry)
XX DE Rat CARL (Cardiac related Ankyrin-Repeat Protein Kinase).
XX Cardiac related ankyrin repeat protein kinase; CARL; cytoskeleton;
XX cardiac cell growth factor receptor; cell differentiation; modulator;
XX regulator; detection; cellular proliferation; cardiovascular disorder;
XX heart failure; hypertension; cancer; sarcoma.
XX OS Rattus norvegicus.
XX PN WO200034330-A1.
XX PD 15-JUN-2000.
XX PF 10-DEC-1999; 99WO-US029465.
XX PR 11-DEC-1998; 98US-011938P.
XX PR 14-APR-1999; 99US-00291839.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Raju J;
XX WIPI; 2000-431275/37.
XX DR N-PSDB; AAA47608, AAA47609.
XX New polynucleotide encoding cardiac-related ankyrin-repeat protein
PT kinase, useful for treating disorders such as cardiovascular disorders,
PT e.g. heart failure and cell differentiation disorders, e.g. cancer.
XX Claim 1; Fig 5; 161pp; English.
XX

CC CARK polypeptides are regulators of signal transmission from cellular
CC receptors (e.g. cardiac cell growth factor receptor) and cytoskeleton
CC function. They also act as modulators of cellular differentiation and
CC cell death. Nucleic acids, or their fragments encoding CARK polypeptides
CC are useful for detecting CARK nucleic acids especially mRNA, in a sample.
CC CARK is useful for treating disorders associated with upregulation or
CC downregulation of cellular proliferation such as, cardiovascular
CC disorders (heart failure and hypertension) and disorders associated with
CC cell differentiation such as cancer and sarcoma
XX
SQ

Sequence 835 AA;

Query Match 92.9%; Score 4079; DB 3; Length 835;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 763; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

QY 1 MGNVSRPTOTCTDEWKKVSESIVTIERLEDDLOIKKELTELNIIGSDFAFSKVN 60
DB 1 MGNVSRPTOTCTDEWKKVSESIVTIERLEDDLOIKKELTELNIIGSDFAFSKVN 60
QY 61 NYRTENGSLHLCCICGGKSHIRTLMLKGLRPSRLTRNGFTALHLAVYKDAEITSL 120
DB 61 NYRTENGSLHLCCICGGKSHIRTLMLKGLRPSRLTRNGFTALHLAVYKDAEITSL 120
QY 121 LHSAGDIOQVGGITLHIAITIAHLEADVLLOHGANYIODAVFTPLHTAAYYGH 180
DB 121 LHSAGDIOQVGGITLHIAITIAHLEADVLLOHGANYIODAVFTPLHTAAYYGH 180
QY 181 QVTRLLKFGADVNVSGVGRPLHLASAKGFNIKLLMEBSKADVNAQDNEDVPLH 240
DB 181 QVTRLLKFGADVNVSGVGRPLHLASAKGFNIKLLMEBSKADVNAQDNEDVPLH 240
QY 241 FCSRFGHDIKYLQSLDLEQVPHVNIYGTPTPLHACYNGKFEVAKIIOIGTSLSLK 300
DB 241 FCSRFGHDIKYLQSLDLEQVPHVNIYGTPTPLHACYNGKFEVAKIIOIGTSLSLK 300
QY 301 ENIFSETAFHSACTYKSIDLVKFLLDQNVININHQGRDGTGLHSACVHGHLRVQFL 360
DB 301 ENIFSETAFHSACTYKSIDLVKFLLDQNVININHQGRDGTGLHSACVHGHLRVQFL 360
QY 361 DNGADMLVACDPSRSSEKDEQCLMWAYEKGHDAIVTLLKHYPQDPELPCNEYSQPG 420
DB 361 DNGADMLVACDPSRSSEKDEQCLMWAYEKGHDAIVTLLKHYPQDPELPCNEYSQPG 420
QY 421 GDGSYVSPVPLGKIKSMTEKADILLRAGLPSHFHLOLSEIEFHEIIGSGSGFKVYKG 480
DB 421 GDGSYVSPVPLGKIKSMTEKADILLRAGLPSHFHLOLSEIEFHEIIGSGSGFKVYKG 480
QY 481 RCNRKIVAIKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLDNDPSQFAIVTQ 540
DB 481 RCNRKIVAIKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLDNDPSQFAIVTQ 540
QY 541 YISGSLPSLLHOKRILDLQSLIIADVAKMEYHLNLTQPIIHRDLNSHNLLEYDG 600
DB 541 YISGSLPSLLHOKRILDLQSLIIADVAKMEYHLNLTQPIIHRDLNSHNLLEYDG 600
QY 601 HAVVADFGESRFLQSLDEDNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660
DB 601 HAVVADFGESRFLQSLDEDNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660
QY 661 GEIPFAHLKPAADAAADWYHHRPIIGYIPKPISSLLIRGNWACPEGRPESEVVKLE 720
DB 661 GEIPFAHLKPAADAAADWYHHRPIIGYIPKPISSLLIRGNWACPEGRPESEVVKLE 720
QY 721 ECLCNIELMSPASSNSGSLSPSSSCLNVNRPGRSHVAALRSFELEYALNARSYAG 780
DB 721 ECLCNIELMSPASSNSGSLSPSSSCLNVNRPGRSHVAALRSFELEYALNARSYAG 780
QY 781 LSQAGQYSSOGLSLEEMKRSLOVTPIDKYGVSDPMSMHFSCNRSFEDSS 835
DB 781 WSQSVGTHSNPGLSLEEMNRSTQYSTVDKYGVSDPMSLTHLSRQDDNFEDSN 835

RESULT 6

ABR41887

ID ABR41887 standard; protein; 835 AA.

XX ABR41887;

XX 11-AUG-2003 (first entry)

XX Rat cardiac-related ankyrin-repeat protein kinase (CARK).

XX Rat; CARK; cardiac-related ankyrin-repeat protein kinase; enzyme;

XX Rattus norvegicus.

XX Key Location/Qualifiers

XX Domain 66..99

XX Domain /note= "Ankyrin repeat domain"

XX Domain 100..132

XX Domain /note= "Ankyrin repeat domain"

XX Domain 133..165

XX Domain /note= "Ankyrin repeat domain"

XX Domain 168..198

XX Domain /note= "Ankyrin repeat domain"

XX Domain 199..233

XX Domain /note= "Ankyrin repeat domain"

XX Domain 234..264

XX Domain /note= "Ankyrin repeat domain"

XX Domain 269..302

XX Domain /note= "Ankyrin repeat domain"

XX Domain 306..338

XX Domain /note= "Ankyrin repeat domain"

XX Domain 339..371

XX Domain /note= "Ankyrin repeat domain"

XX Region 411..415

XX Domain /note= "LXCXE motif"

XX Domain 463..716

XX Domain /note= "protein kinase domain"

XX WO2003020912-A2.

XX 13-MAR-2003.

XX 04-SEP-2002; 2002WO-US028300.

XX 05-SEP-2001; 2001US-00947199.

XX (MILL-) MILLENNIUM PHARM INC.

XX Raju J;

XX WPI: 2003-290188/28.

XX N-PSDB: ACC48581.

XX Novel isolated cardiac-related ankyrin-repeat protein kinase polypeptide,
XX useful for treating cellular growth related disorders which include
XX cardiovascular disorders and proliferative and/or differentiative
XX disorders.

Claim 14; Fig 5A-D; 158pp; English.

XX The present sequence is the protein sequence of a novel 96 kDa rat
XX cardiac-related ankyrin-repeat protein kinase, designated CARK, which
XX plays a role in signalling pathways associated with cardiac cell growth
XX or differentiation. CARK is a dual specificity protein kinase having
XX specificity for both serine/threonine and tyrosine residues. Human CARK
XX (see ABR41886) expression is increased in ischaemic heart tissue samples,
XX suggesting a role in the regulation of cardiac cell growth and/or
XX differentiation and the pathogenesis of cardiovascular disorders, e.g.
XX congestive heart failure and cardiac hypertrophy. The invention provides
XX human and rat CARK proteins, fusion proteins, antigenic peptides and anti
XX CARK antibodies, and also CARK nucleic acids, antisense molecules,
XX recombinant expression vectors, host cells and transgenic animals. These

are useful as modulating agents for regulating a variety of cellular processes, e.g. cardiac cellular process, for modulating the phosphorylation state of a CARP molecule or one or more proteins involved in cellular growth or differentiation, for modulating cell behaviour or as targets and therapeutic agents controlling cardiac cell proliferation, differentiation, hypertrophy and migration, for modulating intra- or inter-cellular signalling and/or gene transcription, for modulating cell proliferation, growth, differentiation, survival and/or migration, for regulating transmission of signals from cellular receptors, for modulating entry of cells, e.g. cardiac precursor cells, into mitosis, or for regulating cytoskeletal function. The proteins and nucleic acids are useful for treating cellular growth related disorders which include cardiovascular disorders (such as heart failure, hypertension), and proliferative and/or degenerative disorders (such as cancer). They are also useful in screening assays, detection assays (e.g. chromosomal mapping, tissue typing, forensic biology), predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials) and pharmacogenomics, and in methods of treatment. A CARP polypeptide can be used to generate a specific antibody, to screen for naturally-occurring CARP substrates, and to screen for drugs or compounds which modulate CARP activity. It is also useful as a bait protein in a yeast two-hybrid or three-hybrid assay and for identifying other proteins which bind to, or interact with, CARP or which are involved in CARP activity

XX SQ Sequence 835 AA;

Query Match 92.9%; Score 4079; DB 6; Length 835;
 Best Local Similarity 91.4%; Pred. No. 0;
 Matches 763; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

QY 1 MGNYSRPTCTDEWKKVSESVYTIERLEDLQIKEKELTELRNIFGSDPAFSKYNL 60
 DB 1 MGNYSRPTCTDEWKKVSESVYTIERLEDLQIKENEFQELRHIFGSDPAFSEVSL 60

QY 61 NYTENGSLHLHCCTCGGKKSHIRTLMLKGLRPSRLTRNGTALTALHAYKDNELITSL 120
 DB 61 NYRTERGLSLHLCCVCGGKSHIRALMLKGLRPSRLTRNGFPALHAYKDSPELITSL 120

QY 121 LHSAGDIQQVGYGGLTALHATATAGLEAADVLQHGANNVQDAVFTPLHIAAYGHE 180
 DB 121 LHSAGVQVGYGGLTALHATATAGPEAAEVLLQHGANNVQDAVFTPLHIAAYGHE 180

QY 181 QVTRLLKFGADVNSGEVDRPLHLASAKGFLNTAKLMEEGSKADVNAQDNEDHVPFH 240
 DB 181 QVTSVLLKFGADVNSGEVDRPLHLASAKGFNIVKLLVEEGSKADVNAQDNEDHVPFH 240

QY 241 FCSRFGHDIKVKLLQSDLEVOHVNIVYDTPHLACVNGKFEVAKETIQISGESLTK 300
 DB 241 FCSRFGHNIIVSYLLQSDLEVOHVNIVYDTPHLACVNGKFEVAKETIVQTGESLTK 300

QY 301 ENIFSETAFHSACTYGSIDLVKFLLDQNVININHQGRDGTGLHSACVHGHIRLVQELL 360
 DB 301 ENIFSETAFHSACTYGNIDLVKFLLDQNAVININHGRDGTGLHSACVHGHIRLVQELL 360

QY 361 DNGADNMLVACDSRSRSGEKDEOTCLMWAYEKHDAIVTLKHYKRPQDELPCNYSQPG 420
 DB 361 DNGADNMLVACDSRSRSGEKDEOTCLMWAYEKHDAIVTLKHYKRPQDELPCNYSQPG 420

QY 421 GDGSYVSVPSPLGKISMTKEKADILLRAGLPSHFHLSIEPHEIIGSGSFQKVKYG 480
 DB 421 GDGSYVSVPSPLGKISMTKEKADVLLRABLPSPHFHLSIEPHEIIGSGSFQKVKYG 480

QY 481 RCRNKIVAKRYRANTYVCSKSDVDMFCREVSILCQLNHPCVITQVGCALNDFSPQAIVTQ 540
 DB 481 RCRNKIVAKRYRANTYVCSKSDVDMFCREVSILCQLNHPCVITQVGCALNDFSPQAIVTQ 540

QY 541 YISGGSLSLHLHQRKRLDLSKLITAVDVAKGMEYLHNLTPHRLDLSNHLILYEDG 600
 DB 541 YISGGSLSLHLHQRKRLDLSKLITAVDVAKGMEYLHSLTPHRLDLSNHLILYEDG 600

QY 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRTYIKADVPFVSCALCWEILT 660
 DB 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRTYIKADVPFVSCALCWEILT 660

QY 661 GEIPFAHLKPAAAAADMAHYHRRPIGYSPKPISSLLIRGNACPEGRPFSEVVMKLE 720
 DB 661 GEIPFAHLKPAAAAADMAHYHRRPIGYSPKPISSLLIRGNACPEGRPFSEVVMKLE 720

QY 721 ECLCNELMSPASSNSSLSPSSSDCLVNRGGPRSHVAALRSRFEVLEVALNARSAA 780
 DB 721 ECLCNELMSPASSNSSLSPSSSDCLVNRGGPRSHVAALRSRFEVLEVALNARSAG 780

QY 781 LSSAQOYSSQGLSLEEMKRSQYTPIDKYGVSDPMSMHFSCNRSSEFSDSS 835
 DB 781 WSQSVGTHSNPGLSLEEMNRSTQYSTDVKYGVSDPMSLTLHLSRQDDNFEDSN 835

RESULT 7 AAEE37962

ID AAEE37962 standard; protein; 603 AA.

AC AAEE37962;

DT 06-NOV-2003 (first entry)

DE Human kinase and phosphatase (KPP-7) protein.

XX Human; kinase; phosphatase; KPP; cell proliferative disorder; hepatitis;
 XX atherosclerosis; cirrhosis; haemoglobinuria; polycythaemia vera; cancer;
 XX psoriasis; thrombocytopenia; developmental disorder; Reiter's syndrome;
 XX renal tubular acidosis; anaemia; mental retardation; Alzheimer's disease;
 XX neurological disorder; Parkinson's disease; epilepsy; glomerulonephritis;
 XX autoimmune disorder; inflammatory disorder; AIDS; Goodpasture's disease;
 XX acquired immune deficiency syndrome; immunosuppressive; Crohn's disease;
 XX neurotic; transgenic; dermatitis; multiple sclerosis; diabetes mellitus;
 XX allergy; gout; Grave's disease; Hashimoto's thyroiditis; bowel syndrome;
 XX osteoporosis; rheumatoid arthritis; Sjogren's syndrome; ophthalmological;
 XX gene therapy; asthma; anticonvulsant; uropathic; pancreatitis.

XX Homo sapiens.

XX W02003050084-A2.

XX 19-JUN-2003.

XX 06-DEC-2002; 2002WO-US039126.

XX 07-DEC-2001; 2001US-0340235P.

XX 19-DEC-2001; 2001US-0343007P.

XX 21-DEC-2001; 2001US-0343546P.

XX 04-FEB-2002; 2002US-0354388P.

XX 15-FEB-2002; 2002US-0357675P.

XX (INCY-) INCYTE GENOMICS INC.

XX Kable AE, Chien D, Wilson AD, Swarnakar A, Gorvad AE;

XX Hafalia AJA, Emerling BM, Ramkumar J, Jin P, Griffin JA, Marquis JP;

XX Baughn WR, Chawla NK, Lehr-Mason PM, Khare R, Lee S, Hawkins PR;

XX Becha SD, Lee SY, Sprague WW, Zebardjian Y;

XX WPI; 2003-532894/50.

XX N-PSDB; AAD57334.

XX New human kinases and phosphatases and polynucleotides, useful for

XX diagnosing, treating or preventing autoimmune or inflammatory disorders

XX (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,

XX cancer or hepatitis.

XX Claim 1; Page 200-201, 282pp; English.

XX The invention relates to an isolated polypeptide, which is a human kinase

XX and phosphatase (KPP). KPP agonists and antagonists are useful for

XX diagnosing, treating or preventing disorders associated with aberrant

XX expression of KPP, particularly cell proliferative disorders (e.g.

XX arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal

XX nocturnal haemoglobinuria, polycythaemia vera, psoriasis, primary

CC thrombocytopenia or cancer), developmental disorders (eg. renal tubular
CC acidosis, anaemia or mental retardation), neurological disorders (e.g.
CC Alzheimer's disease, Parkinson's disease or epilepsy), autoimmune/
CC inflammatory disorders (e.g. AIDS; acquired immune deficiency syndrome,
CC allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's
CC disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome,
CC gout, Grave's disease, Hashimoto's thyroiditis, irritable bowel syndrome,
CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
CC syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
CC bacterial, fungal, parasitic, protozoan or helminthic infections. The KPP
CC is useful in assessing the effects of exogenous compounds on the
CC expression of nucleic acids and kinases and phosphatases. KPP gene is
CC useful in gene therapy and for creating transgenic animals to model human
CC disease. The present sequence is human KPP protein
XX
SQ Sequence 603 AA;

Query Match 70.7%; Score 3102; DB 6; Length 603;
Best Local Similarity 100.0%; Pred. No. 1e-276;
Matches 590; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGNYSRPTQTCTDEWKKKSESYYTIERLEDLQIKEXELTELRLNIFGSDFAFSKVL 60
DB 1 MGNYSRPTQTCTDEWKKKSESYYTIERLEDLQIKEXELTELRLNIFGSDFAFSKVL 60
QY 61 NYRTENGSLSLHLCICGGKKSHRTLMKGLRPSRLTRNGFTALHLAVYKDAEILTSI 120
DB 61 NYRTENGSLSLHLCICGGKKSHRTLMKGLRPSRLTRNGFTALHLAVYKDAEILTSI 120
QY 121 LHSAGDIOQVGGYGLTALHATIAHLEAADVLLQHGANYNIQDAVFPTPLHIAAYGHE 180
DB 121 LHSAGDIOQVGGYGLTALHATIAHLEAADVLLQHGANYNIQDAVFPTPLHIAAYGHE 180
QY 181 QVTRLLKFGADVNSGVGDRPLHLASAKGFLNIAKLMEBSKADVNAQDNEDHVPILH 240
DB 181 QVTRLLKFGADVNSGVGDRPLHLASAKGFLNIAKLMEBSKADVNAQDNEDHVPILH 240
QY 241 FCSRGHHDIVKYLQSDLEVPQHVNNIYGDTPHLACYNGKEFEVAKIITQSGTESLTK 300
DB 241 FCSRGHHDIVKYLQSDLEVPQHVNNIYGDTPHLACYNGKEFEVAKIITQSGTESLTK 300
QY 301 ENIFSETAFHSACTYGSIDLVKFLDQNVNINHQGRDGHGTGLHSACYHGHIRLVQFLL 360
DB 301 ENIFSETAFHSACTYGSIDLVKFLDQNVNINHQGRDGHGTGLHSACYHGHIRLVQFLL 360
QY 361 DNGADNVUACDPSRSSEKDEQTCMLWAVEKGDHDAIVTLKHYKRPQDELPCNEYSQPG 420
DB 361 DNGADNVUACDPSRSSEKDEQTCMLWAVEKGDHDAIVTLKHYKRPQDELPCNEYSQPG 420
QY 421 GDGSYVSPVPLGKIKMTKEKADILLRAGLPSHFHLQSEIEFHEIIGSGSFGKVKYKG 480
DB 421 GDGSYVSPVPLGKIKMTKEKADILLRAGLPSHFHLQSEIEFHEIIGSGSFGKVKYKG 480
QY 481 RCNRKIVAIKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLNDSQFAIVTQ 540
DB 481 RCNRKIVAIKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLNDSQFAIVTQ 540
QY 541 YISGGSLSLLHCKRILDLQSLKIIAIVDAKMGVEYLHNLTPRIHRDLN 590
DB 541 YISGGSLSLLHCKRILDLQSLKIIAIVDAKMGVEYLHNLTPRIHRDLN 590

RESULT 8
AAU20525
ID AAU20525 standard; protein; 148 AA.
XX
AC AAU20525;
XX
DT 06-DEC-2001 (first entry)
XX
DE Human secreted protein, Seq ID No 517.
XX
KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;

KW rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cystic; Alzheimer's disease; Parkinson's disease; human; cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing.
XX
OS Homo sapiens.
XX
FN WO200155326-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001347.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-020515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-021647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249254P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451931/48.
XX N-PSDB; AAS33234.
XX New nucleic acids and polypeptides, useful for diagnosing, preventing or
XX treating medical conditions.
XX Claim 11; SEQ ID NO 517; 753pp; English.
XX The invention relates to novel isolated nucleic acid molecules (I)
XX encoding human secreted proteins (II). (I) and (II) are used to prevent,
XX treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
XX goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
XX the prevention, treatment and diagnosis of diseases associated with
XX inappropriate expression of secreted proteins. (I) and complementary
XX sequences may also be used as DNA probes in diagnostic assays (e.g.
XX polymerase chain reactions (PCR)) to detect and quantitate the presence
XX of similar nucleic acid sequences in samples, and so which patients may
XX be in need of restorative therapy. (II) may also be used as antigens in
XX the production of antibodies and in assays to identify modulators
XX (agonists and antagonists) of the expression and activity of the secreted
XX proteins. The anti-(II) antibodies and antagonists may also be used to
XX down regulate expression and activity of (II). The anti-(II) antibodies
XX may also be used as diagnostic agents for detecting the presence of (II)
XX in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The
XX disorders include for example: immune/autoimmune diseases (e.g. HIV
XX (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
XX (human immunodeficiency virus) infections, cancers and hyperproliferative disorders (e.g.
XX melanomas, neoplasms of the breast or liver, Sezary syndrome and
XX Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
XX Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
XX cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
XX thrombosis), infections caused by bacteria, viruses and fungi and ocular
XX disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
XX and antibodies can also be used to promote wound healing, maintain organs
XX before transplantation, and support cell culture of primary tissues.

Query Match 13.7%; Score 603; DB 4; Length 148;

Best Local Similarity 96.5%; Pred. No. 5.4e-47;

Matches 109; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 304 FSETAFHSACTYCKSIDLVKFLDQNVININHGQRDGHTGLHSACHGHIRLVQFLDNG 363
| : |||||
Db 15 FMKQFHSACTYCKSIDLVKFLDQNVININHGQRDGHTGLHSACHGHIRLVQFLDNG 74

Qy 364 ADMNLVACDPSRSSGKBDQTCMLWAYEKGHDAIVTLLKHYPQDELPFCNEY 416
| : |||||
Db 75 ADMNLVACDPSRSSGKBDQTCMLWAYEKGHDAIVTLLKHYPQDELPFCNEY 127

RESULT 9

AAU20648

ID AAU20648 standard; protein; 148 AA.

XX AAU20648;

XX AC AAU20648;

XX 06-DEC-2001 (first entry)

XX Human secreted protein, Seq ID No 640.

XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
XX rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;
XX cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
XX cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;

KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing.
XX Homo sapiens.
OS
XX WO200155326-A2.
PN
XX
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001347.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240560P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 01-NOV-2000; 2000US-0244674P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.


```
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149823P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.

PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.2%; Score 493; DB 3; Length 589;
Best Local Similarity 31.6%; Pred. No. 7.6e-36;
Matches 136; Conservative 86; Mismatches 145; Indels 64; Gaps 15;

QY 416 YSOPGGDSYVSPSLGKIKSM--TKKA-----DILLRAGLPS---HFHL 458
DB 51 WSQETDGLRDALSKELKLDQPGSKQKSIFFEHDKSSNELIPACIPIPTDGTDEWEI 110
QY 459 QLSIEFHEITIGSGFGKVKYGRNKIIVAKRYRANTYCSKSDVDMPCREVSILCOLNH 518
DB 111 DVTQKIEKKVAGSGYGLHRGTYSQVEAIRFLKPD--VNNEMLEFESQEVFIRKVRH 169
QY 519 PCVIOFVACLNDSQPAIVTVQISGSLFSLHEOKRILDLQSKLIIADVAKGWEYLH 578
DB 170 KNVQFLGACTRSPT-LCIVTFMARGSIYDFLHKQKAPKLOTLKVALDVAKGWEYLH 228
QY 579 NLTOPIIHRDLNSHNLILYEDGHAVVADFGESEFLOSLEDNNTKOPGNLRWMAPEVTQ 638
DB 229 Q--NNIHRDLKTNLLMDEHGLVKVADFGVARV--QIESGVMTAETGYRWMAPEVI-E 283
QY 639 CTRYTIKADYFVALCLWEILTCETIPFAHLKPAADWYHHIRPPICYSIPKPISSIL 698
DB 284 HKPYNHKADYFVAIVLWELLTGDIPYAFITPQAAVGVQVGLRPIKPKTHPKVKGILL 343
QY 699 IRGNACPEGRPEFSEVMKLEBCLNIELMSPASSNSGSLIS-----PSSSSD 747
DB 344 ERCWHQDPEORPLFEETIEMLQOI-----MKEPVTVFGSATAVEEMVFLSWGRFSSEQ 398
QY 748 CLV-----NRGGPRGRSHVAALRSRFELE---YALNARSVAALSQAGQYSSQG 792
DB 399 QQVINKTGTFFNYDNKYRGSRS--IAKLKEDSEIDKXGFLIN---HARVLVSGRE---- 450
QY 793 LSLEMKRSIQ 803
```

```
Db      451  -SYEKGKALQ 460
RESULT 11
AAG45983
ID  AAG45983 standard; protein; 732 AA.
XX
AC  AAG45983;
XX
DT  18-OCT-2000 (first entry)
XX
DE  Arabidopsis thaliana protein fragment SEQ ID NO: 57799.
XX
KW  Protein identification; signal transduction pathway; metabolic pathway;
KW  hybridisation assay; genetic mapping; gene expression control; promoter;
KW  termination sequence.
XX
OS  Arabidopsis thaliana.
XX
PN  EP1033405-A2.
XX
PD  06-SEP-2000.
XX
XX  25-FEB-2000; 2000EP-00301439.
XX
XX  25-FEB-1999; 99US-0121825P.
PR  18-JUN-1999; 99US-0139455P.
PR  18-JUN-1999; 99US-0139456P.
PR  18-JUN-1999; 99US-0139457P.
PR  18-JUN-1999; 99US-0139458P.
PR  18-JUN-1999; 99US-0139459P.
PR  18-JUN-1999; 99US-0139460P.
PR  18-JUN-1999; 99US-0139461P.
PR  18-JUN-1999; 99US-0139462P.
PR  18-JUN-1999; 99US-0139463P.
PR  18-JUN-1999; 99US-0139750P.
PR  18-JUN-1999; 99US-0139763P.
PR  21-JUN-1999; 99US-0139817P.
PR  21-JUN-1999; 99US-0139899P.
PR  23-JUN-1999; 99US-0140353P.
PR  23-JUN-1999; 99US-0140354P.
PR  24-JUN-1999; 99US-0140695P.
PR  28-JUN-1999; 99US-0140823P.
PR  29-JUN-1999; 99US-0140991P.
PR  30-JUN-1999; 99US-0141287P.
PR  01-JUL-1999; 99US-0141842P.
PR  01-JUL-1999; 99US-0142154P.
PR  02-JUL-1999; 99US-0142055P.
PR  06-JUL-1999; 99US-0142390P.
PR  08-JUL-1999; 99US-0142803P.
PR  09-JUL-1999; 99US-0142920P.
PR  12-JUL-1999; 99US-0142977P.
PR  13-JUL-1999; 99US-0143542P.
PR  14-JUL-1999; 99US-0143624P.
PR  15-JUL-1999; 99US-0144005P.
PR  16-JUL-1999; 99US-0144085P.
PR  16-JUL-1999; 99US-0144086P.
PR  19-JUL-1999; 99US-0144325P.
PR  19-JUL-1999; 99US-0144331P.
PR  19-JUL-1999; 99US-0144332P.
PR  19-JUL-1999; 99US-0144333P.
PR  19-JUL-1999; 99US-0144334P.
PR  19-JUL-1999; 99US-0144335P.
PR  20-JUL-1999; 99US-0144352P.
PR  20-JUL-1999; 99US-0144632P.
PR  20-JUL-1999; 99US-0144884P.
PR  21-JUL-1999; 99US-014814P.
PR  21-JUL-1999; 99US-0145086P.
PR  21-JUL-1999; 99US-0145088P.
PR  22-JUL-1999; 99US-0145085P.
PR  22-JUL-1999; 99US-0145087P.
PR  22-JUL-1999; 99US-0145089P.
PR  22-JUL-1999; 99US-0145192P.
PR  23-JUL-1999; 99US-0145145P.
PR  23-JUL-1999; 99US-0145218P.
PR  23-JUL-1999; 99US-0145224P.
PR  26-JUL-1999; 99US-0145276P.
PR  27-JUL-1999; 99US-0145913P.
PR  27-JUL-1999; 99US-0145918P.
PR  27-JUL-1999; 99US-0145919P.
PR  28-JUL-1999; 99US-0145951P.
PR  02-AUG-1999; 99US-0146386P.
PR  02-AUG-1999; 99US-0146388P.
PR  02-AUG-1999; 99US-0146389P.
PR  03-AUG-1999; 99US-0147038P.
PR  04-AUG-1999; 99US-0147204P.
PR  04-AUG-1999; 99US-0147302P.
PR  05-AUG-1999; 99US-0147192P.
PR  05-AUG-1999; 99US-0147260P.
PR  06-AUG-1999; 99US-0147303P.
PR  06-AUG-1999; 99US-0147416P.
PR  09-AUG-1999; 99US-0147493P.
PR  09-AUG-1999; 99US-0147935P.
PR  10-AUG-1999; 99US-0148171P.
PR  11-AUG-1999; 99US-0148319P.
PR  12-AUG-1999; 99US-0148341P.
PR  13-AUG-1999; 99US-0148565P.
PR  13-AUG-1999; 99US-0148684P.
PR  16-AUG-1999; 99US-0149368P.
```

```
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149732P.
PR 20-AUG-1999; 99US-0149733P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151085P.
PR 27-AUG-1999; 99US-0151086P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158023P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162143P.

Query Match 11.2%; Score 493; DB 3; Length 732;
Best Local Similarity 31.6%; Pred. No. 1.1e-35;
Matches 136; Conservative 86; Mismatches 145; Indels 64; Gaps 15;

QY 416 YSQPGDGSYVSPFLGKIKSM--TKEA-----DILLRAGLPS---HFHL 458
DB 194 WSQETDGLRDLKXELKLDQPGSKQKISFPEHDKSNELIPACIEIPTDGTDEWEI 253
QY 459 QLSIEFHEIIIGSGFGKYKGRCKNIKIVAKRYRANTYKSKSDVMFCREVSIQLCOLNH 518
DB 254 DVTQLKIERKVASGSGDLHRTYCSQEVAIKFLKPDPR-VNNEMLREFSQEVFTMRKVRH 312
```

```
QY 519 PCVIQFVAGACLNDPQFAIVTQVYISGSLFSLHBEQKRIIDLOSKLIITAVDAKMEYLH 578
DB 313 KNVVQFLGACTRSPT-LCIVTFPMARGSIYDFLHKOKCAFKLQTLTKVALDVAKGMSYLH 371
QY 579 NLTOPIIHRDLNHNILLYEDGHAVVADFGESRFLQSLDEDNNTKQPGNLRWMAPEVTO 638
DB 372 Q--NNIIHRDLKTANLMDHGLVKVADFGVARV--QIESGVMTAETGTYRWMAPEVI-E 426
QY 639 CTRYTIKADYFSYALCLWEILTGEIPFAHLKPAAAAADMAHYHHRPPIGYSIPKISSILL 698
DB 427 HKPYNHKADYFSAIVLWELLTGDIPYAFLTPLQAAVGVVQGLRKPDKPKTHPKVKGILL 486
QY 699 IRGNWACPEGRPEFSEVWKLECLCNIELMSPASSNSSGSLs-----PSSSSD 747
DB 487 ERCWHQDPEQRPLFEEIEMLOOI-----MKEPVTVFGSASTAIVEMVFLSWGRFSSEQ 541
QY 748 CLV-----NRGGPGRSHVAALRSRFELE---YALNARSYAALSQAGYSSQG 792
DB 542 QQVINKTGTENYDNKYGVSSRS-IAKIKEDSEIDKDGFLIN---HARVLVSGRE---- 593
QY 793 LSLEEMKRSIQ 803
DB 594 -SYEKGKKALQ 603

RESULT 12
AAG45982
ID AAG45982 standard; protein; 760 AA.
XX
AC AAG45982;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 57798.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127452P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
```

PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139450P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145244P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147018P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147322P.
PR 05-AUG-1999; 99US-0147323P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151320P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 16-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154019P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.

```
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161362P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161921P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161923P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.2%; Score 493; DB 3; Length 760;
Best Local Similarity 31.6%; Pred. No. 1.2e-35;
Matches 136; Conservative 86; Mismatches 145; Indels 64; Gaps 15;

QY 416 YSQPGDGSYVSPGKIKSM--TKEKA-----DILLRAGLPS-----HFHL 458
DB 222 WQETDGLRDALEKILKQDPGSKQKSIPEHDKSNELIPACIEIPTDGTDEWEI 281
QY 459 QLSIEFHEIIGSGFGKVKYGRCKNIVAKRYRANTYCSKSDVDMFCREVSILCOLNH 518
DB 282 DVTQLKIEKVASGSGDLHREGTYSQSEVAIKFLKPRD-VNNEMLREFSOEVEFMRKVRH 340
QY 519 PCVIOFVACLNDSPFAIVQYISGSLFSLLEHQRILDSKLIIVADVAKMEYLH 578
DB 341 KVVQFLGACTRSPT-LCIVTFVARGSIYDFLHKQKCAFLOTLLKVALDVAKMSYLH 399
QY 579 NLQPTIHRDLNSHNLIELYDEGHAVADFGESRFLQSLDDEMTKQGNLRWMAPEVFTQ 638
DB 400 Q--NNIIHRDLKTNALLMDEHGLVKVADFGVARY--QIESGVMTAETGTYYRWMAPEVI-E 454
QY 639 CTRYTIKADVSFALCWEILITGTETPAHLKPAAPAAADMAVYHRRPPIGYSIKPISL 698
DB 455 HKPYNHKAQVSFALVWELLTGDIPTAFLTPLOAAVGVQVKGLRKPDKTTPKVKGLL 514
QY 699 IRGNWACPEGRPEFSEVVMKLEECNLIELMSPASSNSGSL-----PSSSD 747
DB 515 ERCWQDPEQPLPEEIIEMLQI-----MKEPTVFGSASIAVEEMVFLSGRPSSEQ 569
QY 748 CLV-----NRGPGRSVHAALRSRFELE---YALNARSVAALQSAGQYSSQG 792
DB 570 QQVINKTGTFTYNDKRYGVSSRS-TAKLEDSEIDKDGFLIN---HARVLVSGRE---- 621
QY 793 LSLSEEMKSLQ 803
DB 622 -SYEGKALQ 631

RESULT 13
ADD47763
ID ADD47763 standard; protein; 1881 AA.
XX AC ADD47763;
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein XP_016774, SEQ ID NO 13459.
XX KW Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GEO ) GEN HOSPITAL CORP.

(PARB ) BAYER AG.
Woelf C, D'urso D, Befort K, Costigan M;
WPI; 2003-268312/26.
GENBANK; XF_016774.
New composition comprising two or more isolated polypeptides, useful for
preparing a medicament for treating pain in an animal.
Claim 1; Page; 1017pp; English.
The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition, a
method for identifying a compound or small molecule that regulates the
activity in an animal of one or more of the polypeptides given in the
specification, a method for identifying a compound useful in treating
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (SNI), chronic constriction
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
therapy). The sequence presented is a human protein (shown in Table 2 of
the specification) which is differentially expressed during pain. Note:
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic form directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.
Sequence 1881 AA;
Query Match 10.5%; Score 460.5; DB 7; Length 1881;
Best Local Similarity 27.6%; Pred. No. 5.2e-32;
Matches 166; Conservative 76; Mismatches 176; Indels 183; Gaps 23;
QY 25 VITIERLED---DLOIKEK-ELTEL-----RNIFGSDEAFSKYNLNY-----RTENGLSL 70
DB 252 VIMVRLLDGRGAQIETKTDELTPHCAARN---GHRVRISEILLDHGAPIQAKTNGULSP 308
QY 71 LHLG-----CI-----CGGKKSHIRTMILKGLRPSRL 97
DB 309 IHMAAQGDHLDVRLQLQYDAEIDITLDHLTPHVAACHGHRV-AKVLDDKGAKNR 367
QY 98 TRNGFTALHAYKDNALITSLHSGADIQQVGYGLTALHTATTAGHLEADVLQHG 157
DB 368 ALNGFTPLHIAACKNHYRVNMLLKTGASIDATESGLTPHVASFMGHLPIVKLLQRG 427
QY 158 ANVNIQDAVFPTPLHIAAYYGHEQVTRLLIKFGADVNVSGVEGDRPLHLASARGFLNTAK 217
DB 428 ASPNVSNKVTETPLHMAARAGHTEVAKYLLQNKAKVNAKAKDDQDTLHCAARIGHTNMVK 487
QY 218 LLMEEGSKADVNAODNEDHVLPHFCRFGH-----HDIVK 252
DB 488 LLLS--NNANPNLATTAGHTPLHIAAREGHVETVLALLEKEASQACMCKKGTPLHLVAAK 545
QY 253 Y-----LLOSD-----LEV-----OPHVNYIGDTP 273
DB 546 YGKRVVAELLERDAHPNAAGKNGLTPHVAHVHNNLDIVKLLPRGSGHSPAMNGYTP 605
QY 274 LHLACYNGKFEVAKIITIGTSESLTKENIFSTAFHSACTYKGSIDLKVLDDQNV--- 330
DB 606 LHIAAKQNVQEVARSLLQYGG--SANAESVQGVTPFLHAAQEGHA-EMVALLLSKOANGN 662
```

QY 331 -----ININHQGDGHTGLHSACYHGHIRLVQFLLD 361
 Db 663 LGNKSGLTPLHLVAQEGHVPADVLIKHGMVMDATTRMGYTPLHVASHYGN:KLKVKFLQ 722
 QY 362 NGADMLVACDPSRSSGEKDEQCLMWAYEKGHDAIVT--LLKHVKRPQDELFCNEYSQPG 420
 Db 723 HQADV-----AKTKLGYSPLHQAAQOQHTDIVTLKNGASP-----NEVSDG 767
 QY 421 GDG-----SYVSVSPGLKIKSMTEKADILL---RAGLPSPHFH--LQLSETEFHEI 468
 Db 768 TTPLAIARLGIYSVDVL---KVVTDTSFVLVSDKHRMSPFETVDEILDVSEDEGEL 824
 QY 469 I 469
 Db 825 I 825

RESULT 14
 ABU11523
 ID ABU11523 standard; protein; 1536 AA.
 AC ABU11523;
 XX
 XX 12-FEB-2003 (first entry)
 DT Human MDDT polypeptide SEQ ID 470.
 XX
 DE MDDT; human; disease detection and treatment molecule polypeptide;
 KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;
 KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;
 KW gene therapy; protein replacement therapy; cell proliferative disorder;
 KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;
 KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
 KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;
 KW psoriasis; hepatitis.
 XX
 OS Homo sapiens.
 XX
 XX W0200279449-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 27-MAR-2002; 2002WO-US009944.
 XX
 PR 28-MAR-2001; 2001US-0279619P.
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-0299776P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 PI Dufour GS, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefteld Y, Gerstin RH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX
 XX WPI; 2003-058431/05.
 DR N-PSDB; ABX34513.
 XX
 XX New purified disease detection and treatment molecule proteins and
 PT polynucleotides, useful for diagnosing, treating or preventing cancers
 PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis
 PT or hepatitis.
 XX
 XX Claim 27; SEQ ID NO 470; 339pp + Sequence Listing; English.
 XX
 XX

CC This invention describes a novel disease detection and treatment molecule
 CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,
 CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,
 CC antianaemic, antipsoriatic and hepatotropic activity. The polynucleotides
 CC and the polypeptides of the invention can be used for gene therapy,
 CC protein replacement therapy and are useful for treating a variety of
 CC diseases or conditions. These polypeptides or polynucleotides are
 CC particularly useful for diagnosing, treating or preventing cell
 CC proliferative disorders (e.g. cancers including adenocarcinoma,
 CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's
 CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's
 CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or
 CC hepatitis. ABU11450-ABU11845 represent the MDDT polynucleotides encoded
 CC by ABU11450-ABU11845, described in the disclosure of the invention. NOTE:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1536 AA;

Query Match 10.3%; Score 454; DB 6; Length 1536;
 Best Local Similarity 28.3%; Pred. No. 1.5e-31;
 Matches 151; Conservative 70; Mismatches 175; Indels 138; Gaps 17;
 QY 45 LRNIFGSDEAFSKVNLNRYRTENGLSLHLCCICGGKXSHIRTLMLKGLRPSRLTRNGFTA 104
 Db 48 VRLLOYDAEIDDTLDH-----LTPHVAHCGHHRV-AKVLLOKGAKPNSRALNGFTF 101
 QY 105 LHVAVKDANALITSLHSGADIQQGYGGLTALHIAIAGHLEAADVLLQGANVITQD 164
 Db 102 LHTACKNHRVNRMLLTKGASIDAVTESGLTPLHVSFMGLPTVKLLQKGASPNYSN 161
 QY 165 AVFTPLHIAAYVGHBOVTRLLKFGADVNSGEGVDRPLHLASAKGFINTAKLMEERGS 224
 Db 162 VKVETPLHVAARAGHTEVAKYLLQNKAKVNAKDDQTPHCAARIGHTNMVKLLJE--N 219
 QY 225 KADVNAQDNEDHVPHFCSRFGH-----HDIVKY----- 253
 Db 220 NNPNLATTAGHTPLHIAAREGHVETVLALLEKEASQACMTKKGTPLHVAAYKGVKVA 279
 QY 254 --LQSD-----LEV-----QPHVVNIYGDTPHLACYN 280
 Db 280 ELLERDAHPNAGKNGLTPLHVAVHNNLDIVKLLFRGSGFHPSPAWNGYTPHIAAQ 339
 QY 281 GKFEVAKIIOISGTESLTKENIFSETAFHSACTYKSIDLVKFLLDQNV----- 330
 Db 340 NQVEVAHSLLOYG--SANAESVQGVTPHLAAQEGHA-EMVALLSKQANGNLGNKSG 396
 QY 331 -----ININHQGDGHTGLHSACYHGHIRLVQFLLDGADNML 368
 Db 397 TPLHLVAQEGHVPADVLIKHGMVMDATTRMGYTPLHVASHYGN:KLKVKFLQADV 455
 QY 369 VACDPSRSSGEKDEQCLMWAYEKGHDAIVT--LLKHVKRPQDELFCNEYSQPGDG--- 423
 Db 456 -----AKTKLGYSPLHQAAQOQHTDIVTLKNGASP-----NEVSDGTTPLAIA 501
 QY 424 ---SYVSVSPGLKIKSMTEKADILL---RAGLPSPHFH--LQLSETEFHEI 469
 Db 502 KRLGIYSVDVL---KVVTDTSFVLVSDKHRMSPFETVDEILDVSEDEGELI 552

RESULT 15
 ABB67412
 ID ABB67412 standard; protein; 1549 AA.
 XX
 XX AC ABB67412;
 XX
 XX DT 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster polypeptide SEQ ID NO 29028.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW

XX	Drosophila melanogaster.
OS	
XX	WO200171042-A2.
XX	
FN	
XX	27-SEP-2001.
PD	
XX	
XX	23-MAR-2001; 2001WO-US009231.
PF	
XX	
XX	23-MAR-2000; 2000US-0191637P.
PR	
XX	11-JUL-2000; 2000US-00614150.
XX	
XX	(PEKE) PE CORP NY.
PA	
XX	
XX	Venter JC, Adams M, Li PWD, Myers EW;
PI	
XX	
XX	WPI; 2001-656860/75.
DR	
DR	N-PSDB; ABL11515.
XX	
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signaling and cell-cell
PT	interactions.
XX	
XX	Disclosure; SEQ ID NO 29028; 21pp + Sequence Listing; English.
PS	
XX	
XX	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC	sequences (AB201840-AB116175) and the encoded proteins (ABB55773-
CC	ABB572072). The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 1549 AA;
XX	
XX	

442	Qy	KADILLRAG----	LPSEFHQLQSEIPEHIIIGSGFGKVKGRCKNIYA-IKRYR	493
839	Db	LMQETLLSDSDSDSCDDLLDHNH-----	YKNATDDELKANYGQDO	878
494	Qy	ANTYCSKSDVDMFCREVSILCOLNHPCVIQFVGACL-----	NDPSQFAIVTQYISGGSL	547
879	Db	KMPDTNTDHDLD--TDVSV-----	LKKELLPNMESCIELTEIGHKPDNVVIARSOVHLGFL	933
548	Qy	FSLLHEQKRILDLQSKLIIVDVAKG--MEYLHNLTQPII-----	HR	587
934	Db	VSEFL-----	VDAGGSMRGYRHNGVRIIVPKACAEPTRICTRYVKPQR	977
588	Qy	DUNSNHILYEDGHAVADFGESFLOSLDEDNMTKP-----	GNLR	629
978	Db	VNPPTPLM---EGEALV-----	SRAILEMSVDGMFLSPITLVEPHVYGTLL	1019

Search completed: September 2, 2004, 14:21:15
Job time : 151 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 2, 2004, 14:18:48 ; Search time 34 Seconds
(without alignments)
1267.873 Million cell updates/sec

Title: US-10-626-173-2

Perfect score: 4390
Sequence: 1 MGNYSRPTQCTDEWKKV.....PMSSMHFHSRNSSEFSS 835

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5S COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCITUS COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4390	100.0	835	3	US-09-291-839-2
2	4390	100.0	835	4	US-09-458-457-2
3	4390	100.0	835	4	US-09-947-199A-2
4	4079	92.9	835	4	US-09-458-457-8
5	4079	92.9	835	4	US-09-947-199A-8
6	465	10.6	1584	4	US-09-457-040B-27
7	435	9.9	1839	2	US-09-172-977-4
8	435	9.9	1839	4	US-09-404-108-4
9	433.5	9.9	994	4	US-10-164-595-38
10	427.5	9.7	1088	3	US-09-082-059-2
11	419	9.5	843	2	US-09-172-977-3
12	419	9.5	843	4	US-09-404-108-3
13	418	9.5	579	4	US-09-529-279-4
14	418	9.5	579	4	US-10-158-895-4
15	418	9.5	579	4	US-09-529-279-15
16	418	9.5	590	4	US-10-158-895-15
17	404.5	9.2	821	1	US-07-928-464-2
18	404.5	9.2	821	1	US-08-003-311B-2
19	404.5	9.2	821	1	US-08-261-432-2
20	404.5	9.2	821	5	PCT-US93-07347-2
21	400.5	9.1	452	3	US-09-035-706-2
22	400.5	9.1	452	3	US-08-955-841-2
23	400.5	9.1	452	4	US-09-390-425-2
24	400.5	9.1	452	4	US-09-566-906-2
25	397	9.0	1745	2	US-09-031-485-33
26	397	9.0	1745	2	US-08-847-429A-33
27	397	9.0	1745	3	US-09-065-474-33

Query Match 100.0%; Score 4390; DB 3; Length 835;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGNYSRPTQCTDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGDEAFSKVNL 60
Db 1 MGNYSRPTQCTDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGDEAFSKVNL 60
Qy 61 NYKTENGLSLHLCCITCGGKKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDNALITSL 120
Db 61 NYKTENGLSLHLCCITCGGKKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDNALITSL 120
Qy 121 LHSGADIQGVGGTALTALHIAIAGHLEAADVLLQHGANVNIQDAVFTPLHIAAYGHE 180
Db 121 LHSGADIQGVGGTALTALHIAIAGHLEAADVLLQHGANVNIQDAVFTPLHIAAYGHE 180
Qy 181 QVTRLLKFGADVNSGVEGDRPLHLASAKGFLNTAKLMBEGSKADVNAQDNEDHVLPH 240
Db 181 QVTRLLKFGADVNSGVEGDRPLHLASAKGFLNTAKLMBEGSKADVNAQDNEDHVLPH 240
Qy 241 FCSRFGHHDIKYLQSDLEVPVNVYIGDTPLHLACVNGKFEVAKELIISGTESTLK 300
Db 241 FCSRFGHHDIKYLQSDLEVPVNVYIGDTPLHLACVNGKFEVAKELIISGTESTLK 300
Qy 301 ENIFSETAFHSACTYKSIDLVKFLDQNVININHQGRDGHGTGLHSACYHGHIRLVQFLL 360
Db 301 ENIFSETAFHSACTYKSIDLVKFLDQNVININHQGRDGHGTGLHSACYHGHIRLVQFLL 360
Qy 361 DNGADNVLVACPSRSSGKDEQTCIMWAYEKGDALVTLLKHYKRPQDELPCEYISQFG 420
Db 361 DNGADNVLVACPSRSSGKDEQTCIMWAYEKGDALVTLLKHYKRPQDELPCEYISQFG 420

RESULT 1
US-09-291-839-2
; Sequence 2, Application US/09291839A
; Patent No. 6261818
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: WNI-068
; CURRENT APPLICATION NUMBER: US/09/291,839A
; CURRENT FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-291-839-2

ALIGNMENTS

28 397 9.0 1745 4 US-09-557-034-33 Sequence 33, Appl
29 389 8.9 263 3 US-09-035-706-5 Sequence 5, Appl
30 389 8.9 263 3 US-08-955-841-5 Sequence 5, Appl
31 389 8.9 263 4 US-09-390-425-5 Sequence 5, Appl
32 389 8.9 263 4 US-09-566-906-5 Sequence 5, Appl
33 382.5 8.7 352 3 US-09-065-474-139 Sequence 139, App
34 382.5 8.7 352 4 US-09-557-034-139 Sequence 139, App
35 377 8.6 455 3 US-09-221-235-5 Sequence 5, Appl
36 377 8.6 455 3 US-09-221-928-5 Sequence 5, Appl
37 377 8.6 455 3 US-09-221-527-5 Sequence 5, Appl
38 377 8.6 455 3 US-09-221-236-5 Sequence 5, Appl
39 377 8.6 455 3 US-09-221-416-5 Sequence 5, Appl
40 377 8.6 455 3 US-09-221-245-5 Sequence 5, Appl
41 377 8.6 455 3 US-09-163-115-5 Sequence 5, Appl
42 377 8.6 455 3 US-09-221-528-5 Sequence 5, Appl
43 377 8.6 455 3 US-09-593-553-5 Sequence 5, Appl
44 377 8.6 455 3 US-09-221-237-5 Sequence 5, Appl
45 377 8.6 455 4 US-09-399-588-2 Sequence 2, Appl

QY	421	GDGSYVSPSLGKIKSMTEKADILLRAGLPSHFLQLSIEFHEIIGSGFGKYKG	480	Db	241	FCSRFGHHDIVKYLQSDLEQPHVNVYGDTPHLACYNKGFEVAKELIIQISGTESTLK	300
Db	421	GDGSYVSPSLGKIKSMTEKADILLRAGLPSHFLQLSIEFHEIIGSGFGKYKG	480	QY	301	ENIFSETAFSACTYCKSIDLVKFLDQNVNINHGDRGHTGLHSACVHGHIRLVQFLL	360
QY	481	RCRNKIVAIKRYRANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLNDRSPQFAIVTQ	540	Db	301	ENIFSETAFSACTYCKSIDLVKFLDQNVNINHGDRGHTGLHSACVHGHIRLVQFLL	360
Db	481	RCRNKIVAIKRYRANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLNDRSPQFAIVTQ	540	QY	361	DNGADNMLVACDPSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPKQDELPCNYSQPG	420
QY	541	YISGSLFSLHHEQKRIIDLOSKLIIADVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG	600	Db	361	DNGADNMLVACDPSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPKQDELPCNYSQPG	420
Db	541	YISGSLFSLHHEQKRIIDLOSKLIIADVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG	600	QY	421	GDGSYVSPSLGKIKSMTEKADILLRAGLPSHFLQLSIEFHEIIGSGFGKYKG	480
QY	601	HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT	660	Db	421	GDGSYVSPSLGKIKSMTEKADILLRAGLPSHFLQLSIEFHEIIGSGFGKYKG	480
Db	601	HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT	660	QY	481	RCRNKIVAIKRYRANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLNDRSPQFAIVTQ	540
QY	661	GEIPPAHLKPAADAAADWAYHRIPIGYSIPKPISSLLIRGNWACPEGRPEFSEVVMKLE	720	Db	481	RCRNKIVAIKRYRANTYCSKSDVDMFCREVSIICQLNHPCVIOFVGACLNDRSPQFAIVTQ	540
Db	661	GEIPPAHLKPAADAAADWAYHRIPIGYSIPKPISSLLIRGNWACPEGRPEFSEVVMKLE	720	QY	541	YISGSLFSLHHEQKRIIDLOSKLIIADVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG	600
QY	721	ECLNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA	780	Db	541	YISGSLFSLHHEQKRIIDLOSKLIIADVAKGMEYLHNLTQPIIHRDLNSHNILLYEDG	600
Db	721	ECLNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA	780	QY	601	HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT	660
QY	781	LSQSAGQYSSQGLSLEEMKRSLOQTPIDKYGVSDPMSMHFSCRNSSSFEDSS	835	Db	601	HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT	660
Db	781	LSQSAGQYSSQGLSLEEMKRSLOQTPIDKYGVSDPMSMHFSCRNSSSFEDSS	835	QY	661	GEIPFAHLKPAADAAADWAYHRIPIGYSIPKPISSLLIRGNWACPEGRPEFSEVVMKLE	720
RESULT 2							
US-09-458-457-2							
; Sequence 2, Application US/09458457							
; Patent No. 6500654							
; GENERAL INFORMATION:							
; APPLICANT: Raju, Jeyaseelan							
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES							
; FILE REFERENCE: mni-068cp							
; CURRENT APPLICATION NUMBER: US/09/458.457							
; EARLIER FILING DATE: 1999-12-10							
; EARLIER FILING DATE: 1998-12-11							
; EARLIER FILING DATE: 1999-04-14							
; NUMBER OF SEQ ID NOS: 9							
; SOFTWARE: PatentIn Ver. 2.0							
; SEQ ID NO 2							
; LENGTH: 835							
; TYPE: PRT							
; ORGANISM: Homo sapiens							
US-09-458-457-2							
Query Match 100.0%; Score 4390; DB 4; Length 835;							
Best Local Similarity 100.0%; Pred. No. 0;							
Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
QY	1	MGNYSRPTQCTCTDEWKKKVSESVITIERLEDDLOIKEKELTELNRNIFGSDEAFSKVNL	60	Db	1	MGNYSRPTQCTCTDEWKKKVSESVITIERLEDDLOIKEKELTELNRNIFGSDEAFSKVNL	60
Db	1	MGNYSRPTQCTCTDEWKKKVSESVITIERLEDDLOIKEKELTELNRNIFGSDEAFSKVNL	60	QY	61	NYRTENGLSLLHLCICCGKSKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDNALITSL	120
QY	61	NYRTENGLSLLHLCICCGKSKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDNALITSL	120	Db	61	NYRTENGLSLLHLCICCGKSKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDNALITSL	120
Db	61	NYRTENGLSLLHLCICCGKSKSHIRTLMLKGRPSRLTRNGFTALHLAVYKDNALITSL	120	QY	121	LHSGADIQVGGTALHIATIAHLEAADVLLQHGANNVIOAVFTPLHIAAYGHE	180
QY	121	LHSGADIQVGGTALHIATIAHLEAADVLLQHGANNVIOAVFTPLHIAAYGHE	180	Db	121	LHSGADIQVGGTALHIATIAHLEAADVLLQHGANNVIOAVFTPLHIAAYGHE	180
Db	121	LHSGADIQVGGTALHIATIAHLEAADVLLQHGANNVIOAVFTPLHIAAYGHE	180	QY	181	QVTRLLKFGADVNSGVDGDRPLHLASAKGFLNIAKLLMEEGSKADVNAQDNEDHVPFLH	240
QY	181	QVTRLLKFGADVNSGVDGDRPLHLASAKGFLNIAKLLMEEGSKADVNAQDNEDHVPFLH	240	Db	181	QVTRLLKFGADVNSGVDGDRPLHLASAKGFLNIAKLLMEEGSKADVNAQDNEDHVPFLH	240
Db	181	QVTRLLKFGADVNSGVDGDRPLHLASAKGFLNIAKLLMEEGSKADVNAQDNEDHVPFLH	240	QY	241	FCSRFGHHDIVKYLQSDLEQPHVNVYGDTPHLACYNKGFEVAKELIIQISGTESTLK	300
QY	241	FCSRFGHHDIVKYLQSDLEQPHVNVYGDTPHLACYNKGFEVAKELIIQISGTESTLK	300	Db	241	FCSRFGHHDIVKYLQSDLEQPHVNVYGDTPHLACYNKGFEVAKELIIQISGTESTLK	300

QY 61 NYRTENGLSLHLCCICGGKKSHIRLMKGLRPSRLTRNGFTALHLAVYKNAELITSL 120
DB 61 NYRTENGLSLHLCCICGGKKSHIRLMKGLRPSRLTRNGFTALHLAVYKNAELITSL 120
QY 121 LHSADIQOQVGYGGLTALHIAATAGHLEAADVLQHGANNVQDAVFFPLHIAAYGHE 180
DB 121 LHSADIQOQVGYGGLTALHIAATAGHLEAADVLQHGANNVQDAVFFPLHIAAYGHE 180
QY 181 QVTRLLKFGADVNVSEVGDRLPLHLASAKGFLNIAKLMEEGSKADVNAQDNEDHVP 240
DB 181 QVTRLLKFGADVNVSEVGDRLPLHLASAKGFLNIAKLMEEGSKADVNAQDNEDHVP 240
QY 241 FCSRFHGHDIIVKLLQSDLEVOHVNNIYDTPHLACVNGKFEVAKETIOISGTESLTK 300
DB 241 FCSRFHGHDIIVKLLQSDLEVOHVNNIYDTPHLACVNGKFEVAKETIOISGTESLTK 300
QY 301 ENIFSETAFHSACTYKSIDLVKFLDQNVNININHQGRDGTGLHSACVHGHIRLVQFLL 360
DB 301 ENIFSETAFHSACTYKSIDLVKFLDQNVNININHQGRDGTGLHSACVHGHIRLVQFLL 360
QY 361 DNGADNMLVACDPSRSGEKDEQTCMLWAEKGHDAIVTLLKHYPQDELPCNEYSPQG 420
DB 361 DNGADNMLVACDPSRSGEKDEQTCMLWAEKGHDAIVTLLKHYPQDELPCNEYSPQG 420
QY 421 GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFLQLSIEFHEIIGSGSGFKVYKG 480
DB 421 GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFLQLSIEFHEIIGSGSGFKVYKG 480
QY 481 RCNKKIVAKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLNDSQFAIVTQ 540
DB 481 RCNKKIVAKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLNDSQFAIVTQ 540
QY 541 YISGGSLSFLHHEOKRILDLQSKLIIAVDVAKGMEYLNHTOPIIHRDLNSHNILLYEDG 600
DB 541 YISGGSLSFLHHEOKRILDLQSKLIIAVDVAKGMEYLNHTOPIIHRDLNSHNILLYEDG 600
QY 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660
DB 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660
QY 661 GEIPFAHLKFAAAADMAVHHIRPPIGYSIPKIPISLLIRGNWACPEGRPEFSEVVMKLE 720
DB 661 GEIPFAHLKFAAAADMAVHHIRPPIGYSIPKIPISLLIRGNWACPEGRPEFSEVVMKLE 720
QY 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780
DB 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780
QY 781 LSQSAQYSSQGLSLEEMKRSLOYTPIDKYGYVSDPMSMHFSCRNSSSPEDSS 835
DB 781 LSQSAQYSSQGLSLEEMKRSLOYTPIDKYGYVSDPMSMHFSCRNSSSPEDSS 835

RESULT 4

US-09-458-457-8

; Sequence 8, Application US/09458457

; Patent No. 6500654

; GENERAL INFORMATION:

; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: mni-068cp

; CURRENT APPLICATION NUMBER: US/09/458,457

; CURRENT FILING DATE: 1999-12-10

; EARLIER APPLICATION NUMBER: 60/111,938

; EARLIER FILING DATE: 1998-12-11

; EARLIER APPLICATION NUMBER: 09/291,839

; EARLIER FILING DATE: 1999-04-14

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: Patent.in Ver. 2.0

; SEQ ID NO 8

; LENGTH: 835

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-458-457-8

Query Match 92.9%; Score 4079; DB 4; Length 835;

Best Local Similarity 91.4%; Pred. No. 0;

Matches 763; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

QY 1 MGNYKSRPTOTCTDEWKXKVSYSYVITIERLEDLQIKEKELTELRNIFGSDAFAKVL 60
DB 1 MGNYKSRPTOTCTDEWKXKVSYSYVITIERLEDLQIKEKELTELRNIFGSDAFAKVL 60
QY 61 NYRTENGLSLHLCCICGGKKSHIRLMKGLRPSRLTRNGFTALHLAVYKNAELITSL 120
DB 61 NYRTENGLSLHLCCICGGKKSHIRLMKGLRPSRLTRNGFTALHLAVYKNAELITSL 120
QY 121 LHSADIQOQVGYGGLTALHIAATAGHLEAADVLQHGANNVQDAVFFPLHIAAYGHE 180
DB 121 LHSADIQOQVGYGGLTALHIAATAGHLEAADVLQHGANNVQDAVFFPLHIAAYGHE 180
QY 181 QVTRLLKFGADVNVSEVGDRLPLHLASAKGFLNIAKLMEEGSKADVNAQDNEDHVP 240
DB 181 QVTRLLKFGADVNVSEVGDRLPLHLASAKGFLNIAKLMEEGSKADVNAQDNEDHVP 240
QY 241 FCSRFHGHDIIVKLLQSDLEVOHVNNIYDTPHLACVNGKFEVAKETIOISGTESLTK 300
DB 241 FCSRFHGHDIIVKLLQSDLEVOHVNNIYDTPHLACVNGKFEVAKETIOISGTESLTK 300
QY 301 ENIFSETAFHSACTYKSIDLVKFLDQNVNININHQGRDGTGLHSACVHGHIRLVQFLL 360
DB 301 ENIFSETAFHSACTYKSIDLVKFLDQNVNININHQGRDGTGLHSACVHGHIRLVQFLL 360
QY 361 DNGADNMLVACDPSRSGEKDEQTCMLWAEKGHDAIVTLLKHYPQDELPCNEYSPQG 420
DB 361 DNGADNMLVACDPSRSGEKDEQTCMLWAEKGHDAIVTLLKHYPQDELPCNEYSPQG 420
QY 421 GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFLQLSIEFHEIIGSGSGFKVYKG 480
DB 421 GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFLQLSIEFHEIIGSGSGFKVYKG 480
QY 481 RCNKKIVAKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLNDSQFAIVTQ 540
DB 481 RCNKKIVAKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLNDSQFAIVTQ 540
QY 541 YISGGSLSFLHHEOKRILDLQSKLIIAVDVAKGMEYLNHTOPIIHRDLNSHNILLYEDG 600
DB 541 YISGGSLSFLHHEOKRILDLQSKLIIAVDVAKGMEYLNHTOPIIHRDLNSHNILLYEDG 600
QY 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660
DB 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT 660
QY 661 GEIPFAHLKFAAAADMAVHHIRPPIGYSIPKIPISLLIRGNWACPEGRPEFSEVVMKLE 720
DB 661 GEIPFAHLKFAAAADMAVHHIRPPIGYSIPKIPISLLIRGNWACPEGRPEFSEVVMKLE 720
QY 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780
DB 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA 780
QY 781 LSQSAQYSSQGLSLEEMKRSLOYTPIDKYGYVSDPMSMHFSCRNSSSPEDSS 835
DB 781 LSQSAQYSSQGLSLEEMKRSLOYTPIDKYGYVSDPMSMHFSCRNSSSPEDSS 835

RESULT 5

US-09-947-199A-8

; Sequence 8, Application US/09947199A

; Patent No. 6660490

; GENERAL INFORMATION:

; APPLICANT: Raju, Jeyaseelan

; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: MNI-068CP2

; CURRENT APPLICATION NUMBER: US/09/947,199A
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: US 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: US 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-947-199A-8

Query Match 92.9%; Score 4079; DB 4; Length 835;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 763; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

QY 1 MGNYSRPTQCTDEWKKKUSEVYIERLEDLOIKEXELTELNIQSGDAPSKVL 60
DB 1 MGNYSRPTQCTDEWKKKUSEVYIERLEDLOIKEXELTELNIQSGDAPSKVL 60

QY 61 NYRTENGSLHLHCICCGKSHIRTLMLKGLRPSRLTRNGFTALHVLAVYKNAELITSL 120
DB 61 NYRTENGSLHLHCICCGKSHIRTLMLKGLRPSRLTRNGFTALHVLAVYKNAELITSL 120

QY 121 LHSAGDIQOVYGGTLALHATIAHLEADAVLLOHGANVIOQAVFFPLHIAAYYGH 180
DB 121 LHSAGDIQOVYGGTLALHATIAHLEADAVLLOHGANVIOQAVFFPLHIAAYYGH 180

QY 181 QVTRLLKFGADVNVSGVGRPLHLASAKGFLIAKLMEBSKADVAQDNEDHVLPH 240
DB 181 QVTRLLKFGADVNVSGVGRPLHLASAKGFLIAKLMEBSKADVAQDNEDHVLPH 240

QY 241 FCSRFGRHDIVKYLQSDLEQVPHVNIYGDTPHLACYNQGEVAKIIOIGTSLSLK 300
DB 241 FCSRFGRHDIVKYLQSDLEQVPHVNIYGDTPHLACYNQGEVAKIIOIGTSLSLK 300

QY 301 ENIFSETAFHSACTYKSIDLVKLLDONVININHOGRDGTGLHSAACHYHRLVQELL 360
DB 301 ENIFSETAFHSACTYKSIDLVKLLDONVININHOGRDGTGLHSAACHYHRLVQELL 360

QY 361 DNGADMNLVADCPSSSGEKDEQTLWAYEKGHDAIVTLKHYKRPQDELPCEYSPQ 420
DB 361 DNGADMNLVADCPSSSGEKDEQTLWAYEKGHDAIVTLKHYKRPQDELPCEYSPQ 420

QY 421 DNGADMNLVADCPSSSGEKDEQTLWAYEKGHDAIVTLKHYKRPQDELPCEYSPQ 480
DB 421 DNGADMNLVADCPSSSGEKDEQTLWAYEKGHDAIVTLKHYKRPQDELPCEYSPQ 480

QY 481 RCNKIIVAKRYRANTYCKSDVDMFCREVSILCOLNHPCVIOFVGACLDNDPQFAIVTQ 540
DB 481 RCNKIIVAKRYRANTYCKSDVDMFCREVSILCOLNHPCVIOFVGACLDNDPQFAIVTQ 540

QY 541 YISGSLFSLHLEQKRIIDLOSKLIIAVDAVKMEYHLNLTQPIIHRDLNLSHLLIYEDG 600
DB 541 YISGSLFSLHLEQKRIIDLOSKLIIAVDAVKMEYHLNLTQPIIHRDLNLSHLLIYEDG 600

QY 601 HAVVADFGESRFLQSLDEDMNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLMBILT 660
DB 601 HAVVADFGESRFLQSLDEDMNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLMBILT 660

QY 661 GEIPFAHLKPAADMAHYHRIPIGYSIPKPISSLLIRGNWACPEGRPEFSEVVMKLE 720
DB 661 GEIPFAHLKPAADMAHYHRIPIGYSIPKPISSLLIRGNWACPEGRPEFSEVVMKLE 720

QY 721 ECLCNELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALSRFELEVALNARSYAA 780
DB 721 ECLCNELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALSRFELEVALNARSYAA 780

QY 781 LQSQAGQYSSQGLSLEEMKRSIQYTPIDKYGVSDFPMSSMHFHSRNSSEFEDSS 835

DB 781 WSQSVGTHSNPGLSLEEMKRSIQYTPIDKYGVSDFPMSLTHLSRQDSDNFEDSN 835

RESULT 6
US-09-457-040B-27
; Sequence 27, Application US/09457040B
; Patent No. 6387641
; GENERAL INFORMATION:
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Bellon, Steve
; TITLE OF INVENTION: Crystallized P38 Complexes
; FILE REFERENCE: VPI/98-14
; CURRENT APPLICATION NUMBER: US/09/457,040B
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 27
; LENGTH: 1584
; TYPE: PRT
; ORGANISM: DICI - Dictyostelium Discoideum
US-09-457-040B-27

Query Match 10.6%; Score 465; DB 4; Length 1584;
Best Local Similarity 37.7%; Pred. No. 1e-35;
Matches 109; Conservative 52; Mismatches 100; Indels 28; Gaps 6;

QY 454 SHFHLQSLSEIEFHEIIGSGSFGKVKRCRNKIVAKRYRANTYCKSDVDMFCREVSIL 513
DB 1280 SEYIEDNFLEFQGTIGKGFGEVKGWRETDAIKIYRQFKTKSLVMFQNEVGIL 1339

QY 514 COLNHPCVIOFVGAC-LNDPSPQFAIVTQISGSL-----PSLLHSEKRIIDLOSKLI 565
DB 1340 SKLRHPNVQFLGACTAGGEDHHCIVTEWVGGSLSRQFLTDHFNLEQNPHI-----RLK 1394

QY 566 IAVDVAKGMEYHLNLTQPIIHRDLNLSHLLI-----YEDGHVAVDFGESRF 612
DB 1395 LALDIAGMNYLHGWTPLIHRDLSSRNILLDHNIDPKNPVSSRQDIKCKISDFGLSR- 1453

QY 613 LQSLDEDMNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLMBILTGEIPFAHLKPA 672
DB 1454 LKKEQASQTSQVGCIPYMAPEVFKGDSN-SEKSDVYSYGMVLFELTSDPEQDMKPMK 1512

QY 673 AADMAHYHRIPIGYSIPKPISSLLIRGNWACPEGRPEFSEVVMKLE 721
DB 1513 MAHLAAVESYRPIPLTSSKWEILTQCWDSNPDSRPTFKQIIVHLKE 1561

RESULT 7
US-09-172-977-4
; Sequence 4, Application US/09172977
; Patent No. 5989863
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/172,977
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1839
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g29491
US-09-172-977-4

Query Match 9.9%; Score 435; DB 2; Length 1839;
Best Local Similarity 27.9%; Pred. No. 1.2e-32;

Db	187	FTPLHVAKYGKLEVANLLQKASPD--AAGKSGLTPLHVAHYDNQKVALLLL--DQG 242
Qy	261	VOPVVNIIYGDTPHLACYNGKFEYAKEIIQI-SGTESLTKENIFSETAFHSACTYKSI 319
Db	243	ASPHAAAKNGYTPPLHIAAKKNQMDIATTLLEVGADANAVTROGIAS---VHLAAQSG-HV 298
Qy	320	DLVKFLLDQV--INNH-----QQRDGHGTGLHSA 347
Db	299	DMVSLLLGLENANVNIISNKSGLTPLHLAAQEDRVNVAEVLVNOGAHVDAQTKNGYTPLHV 358
Qy	348	CYHGHRIVFOILLDNGADMNLVACDPSSRSGEKDEQCLMWAYEKGHDAITVLLKHKXP 407
Db	359	CHYGNIKIVNFIQLHSAKN-----AKTKNGYTPHLQAQAOQHTHIINVLLQNNAS 409
Qy	408	QDELPCEYSPG-----GDGSYVSVSPGLGIKISMTRY 441
Db	410	PNELTWNGTALGIARRLG--YISV---VDTLKIVTEE 442

```

RESULT 11
US-09-172-977-3
  Sequence 3, Application US/09172977
  Patent No. 5989863
  GENERAL INFORMATION:
  APPLICANT: Tang, Y. Tom
  APPLICANT: Guegler, Karl J.
  APPLICANT: Corley, Neil C.
  APPLICANT: Yue, Henry
  TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
  FILE REFERENCE: PF-0615 US
  CURRENT APPLICATION NUMBER: US/09/172,977
  CURRENT FILING DATE: 1998-10-14
  NUMBER OF SEQ ID NOS: 4
  SOFTWARE: PERL Program
  SEQ ID NO 3
  LENGTH: 843
  TYPE: PRT
  ORGANISM: Rattus norvegicus
  FEATURE: -
  OTHER INFORMATION: g1841966
US-09-172-977-3

```

```

; SEQ ID NO 3
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE: -
; OTHER INFORMATION: g1841966
US-09-172-977-3

Query Match          9.5%; Score 419; DB 2; Length 843;
Best Local Similarity 29.2%; Pred. No. 1.2e-31;
Matches 128; Conservative 62; Mismatches 181; Indels 68; Gaps 10

QY      60  LNVRYENGSLSLHLCCICGGKKSHIRITMLKGLRPSRLTRNGFTALHLAVYKDNALITS 119
        : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      257  IDAKTRDGLTPLH-CAARSGHDQVVELLERGAPLLARTKNGLSPLHMAAQGDHVECVKH 315

QY      120  LHSGNADIOQVCGSLTALHATATAGLEAADVLQHGANNVTDQAVFFFTPLHIAAYGH 179
        : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      316  LIQHKAPVDVLTLYLTALHAAHCGHTRVTKLSDRANPNRALNGFFTPLHIACKGNR 375

```

```

463 QY IEFH-----EIIGSQ-----SFGVKYKGRCNKI VAIKRY 492
      :      :||
      :||
381 Db LNAHSDCCRLSSQKYSIVSLFSNEHVL SAGFEIDTPDKFORT-----CLHAAA----- 432
      :||
      :||
493 QY RANTVCSK-----SDVDMCFREVSILCOLNHP-----CVIQFV--GACLNDPSQFAIVT 539
      :||
      :||
433 Db GGNVBCIKLQSSGADPHKKDKCORTPLHVAANCHFHCIETLUTVTGANVNTEDDWGRTA 492
      :||
      :||
540 QY QYISG-----GSLFSLHEQKRIILDQSKLIIVDAVKAGMEY-LHNLTOPIIHR 587
      :||
      :||
493 Db LHVAASDWDRNKITILGNAHNSBELERARELKEK-----EATLCLEFLQLONDANPSI-R 546
      :||
      :||
588 QY DLNNSNILLYEDGHAVVADFCESFIQSL-----DEDNMTKQPCNL-----RWMA 632
      :||
      :||
547 Db DKEGYNSTHY-----AAAYGHRQCLLELLERTNSGFESDSGATKSPHLHAAYNHQA 600
      :||
      :||
633 QY PVUF-----TQCTRTYIK--ADVP----- 649
      :||
      :||
601 Db LEVLLQSLVDLDIRDEKGR TALDLAAPKGHTECVEALINQGASIFVKDNVTKRTPFLHASV 660
      :||
      :||
650 QY -----SYALCILEWIL-TGEIPFA-HLKPAAAAADM-----AYHIRPIGYSIPKP--ISSLLI 699
      :||
      :||
661 Db INGHTLCRLRLLEIADNFEAVDVKDAGQTPMLAVAYGHI-DAVSLLLEKEANVDTVDI 719
      :||
      :||
700 QY RQWNA CPGRPEFSEVWVKLBEICNLIELMSPASSNSSGSLSPSSSDCLVNRGGPGRSH 759
      :||
      :||
720 Db LGCTALHRG-----IMTGHEEC---VQML-----LEQEVILCKDSRG----- 754
      :||
      :||
760 QY VAAALRSRFEYALNARSYAALSQAGYSQGLSLEE--MKRSLOYTPI 807
      :||
      :||
755 Db -----RPLPHYAA--ARGHATW---LSELLQMALSEBDCCKQKQYTPL 794
      :||
      :||

```

```

RESULT 10
US-09-082-059-2
; Sequence 2, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
; APPLICANT: Morrow, Jon S.
; APPLICANT: Devarajan, Prasad
; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identification
; FILE REFERENCE: 44574-5002-US
; CURRENT APPLICATION NUMBER: US/09/082.059A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: 60/047356
; EARLIER FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-082-059-2

```

Qy 432 LGKIKSMTEK---ADIL 446
Db 661 LLSHLAAEEDKVNVDIL 679

RESULT 12
US-09-404-108-3
; Sequence 3, Application US/09404108
; Patent No. 6590077
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN
; FILE REFERENCE: PF-0615 US
; CURRENT APPLICATION NUMBER: US/09/404,108
; CURRENT FILING DATE: 1999-09-23
; EARLIER APPLICATION NUMBER: 09/172,977
; EARLIER FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE: -
; OTHER INFORMATION: g1841966
US-09-404-108-3

Query Match 9.5%; Score 419; DB 4; Length 843;
Best Local Similarity 29.2%; Pred. No. 1.2e-31;
Matches 128; Conservative 62; Mismatches 181; Indels 68; Gaps 10;
Qy 60 LNYRTENGSLHLLCCIGCGKSHITMLKGLRPSRLTRNGFTALHLAVYKDNALITTS 119
Db 257 IDAKTRDGLTPLH-CAARSHDQWELLERGAPELLARTKGLSPLHMAAQGDHVECVKH 315
Qy 120 LLSHGADIQGVGGTALHIAIAGHLEAADVYLLQHGANNVIOAVFTPLHIAAYGH 179
Db 316 LLOHKAPVDVDTLYLTALHVAACHYRVTKLLDKRANPNARALNGFTPLHIAACKNR 375
Qy 180 EQVTRLLKFGADVNSGEVGDPLHLASAKGFLNIKLLMEGSKAD----- 227
Db 376 IKWELLVKGVYQIAITSGLTPIVPAFMGHLNIVLLQLNGASPDVTNIRGETALHM 435
Qy 228 -----VNAODNEDHVLPHFCRFGHHIDIVKYLQSLDLEVPVHVI 268
Db 436 AARAGEVEVVRCLLRNGALVDARAREEQPLHIASRLGKTEIVQLLQH--MAHPDAATT 493
Qy 269 YGDTPLHLACYNGKFEVAKETIIOISGTESTLTKENIFSETAFHSACTYKSIDLVKFLDQ 328
Db 494 NGTPLHISAREGQVDVAVSLLLEAGAAHSLATKGF--TFLHVAAYKG-SLDVAKLLQR 550
Qy 329 NVININHQQRDGTGLHSAHYGHIRLVFLDNGADMNLVACDPSRSSGKDEQTCILMW 388
Db 551 RAA-ADSAGKNGLTPLHVAHYDNQKVALLLLEKASPHATA-----KNGYTPLHI 600
Qy 389 AYKGHDAIVTLKH-----KRPDELPCNEYSQPG-----GGSYVSVSP 431
Db 601 AAKKNQWQIASTLLNNGAETNTVTKGVTPHLASQEGHDMVTLVLEKGANIHMTSKG 660
Qy 432 LGKIKSMTEK---ADIL 446
Db 661 LLSHLAAEEDKVNVDIL 679

RESULT 13
US-09-529-279-4
; Sequence 4, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:

; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-4

Query Match 9.5%; Score 418; DB 4; Length 579;
Best Local Similarity 31.6%; Pred. No. 7.5e-32;
Matches 115; Conservative 67; Mismatches 144; Indels 38; Gaps 13;
Qy 458 LQLSELEFHEIIGSGFGKVKGRCKNKVAIKRYRANTYCKSDVDMCEKREVSILCOLN 517
Db 31 IDYKEIEVEEVGGRGAFGVVCKAKWRKDVAIKQIE-----SESEKAFIVELROLSEVN 85
Qy 518 HPCVIOFVGACLNDRPSQFAIVTYIISGGSFLSLH--EOKRILDLQSKLIIIAVDVAKGME 575
Db 86 HPNIVKLYGACLN---PVCLVMEYAEAGSLYVNLHGAERPLPYTTAAHMSNCLQCSQGA 142
Qy 576 YLNLTOP--IHRDLNSHNIILYDGHAV-VADGESRFLQSLDEDNMTKQPNLRMA 632
Db 143 YLHSM-QPKALHHRDLKPNLLVAGGVTKICDFGTACDIQT----HMTNNKGSAAWMA 197
Qy 633 PEVFTOCTRYTKADVFYSVALCWLWELTGEIPFAHLKPAADAAADWAYHH-IRPPIGYGIP 691
Db 198 PEVP-EGSYNEKCDVFSWGIILWEVITRKPFDIGGPAPRIMWAVENGTRPPLIKNLP 256
Qy 692 KPISLLIRGNWACPEGRPEFSEVYVNMKECL-----CNIELMSPA-----SSNS 737
Db 257 KPISLMTKCSKDPQRSDEEIVKIMTHLMRYFPFGADEPLQYPCQYSDGQSNATST 316
Qy 738 GSLSPSSSDCLVNRGPGSRSHVALRSFEFELEYALNARSYAALSQSAGQYS---SQGLS 794
Db 317 GSPMDIATSN-TSNKSDTNNMQVPATNDTIKRLSKLKNQAKQOSESGRUSLGASHGSS 375
Qy 795 LDEM 798
Db 376 VESL 379

RESULT 14
US-10-158-895-4
; Sequence 4, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 2, 2004, 14:24:19 ; Search time 142 Seconds
(without alignments)

1852.672 Million cell updates/sec

Title: US-10-626-173-2

Perfect score: 4390

Sequence: 1 MGNYSRPTQTCTDEWKVKY.....PMSSMHFHSRNSSEFSDSS 835

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1298764 seqs, 315065143 residues

Total number of hits satisfying chosen parameters: 1298764

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4390	100.0	835	9	US-09-947-199-2
2	4390	100.0	835	16	US-10-626-173-2
3	4079	92.9	835	9	US-09-947-199-8
4	4079	92.9	835	16	US-10-626-173-8
5	1893	43.1	359	16	US-10-664-421-138
6	486.5	11.1	359	16	US-10-437-963-195691
7	483.5	11.0	376	12	US-10-424-599-271998
8	482	11.0	603	16	US-10-437-963-156923
9	480.5	10.9	478	12	US-10-425-114-40404
10	478.5	10.9	585	12	US-10-425-114-67797
11	473	10.8	323	12	US-10-425-114-69530
12	473	10.8	462	12	US-10-425-114-67647
13	472.5	10.8	338	12	US-10-425-114-64160
14	468	10.7	551	12	US-10-424-599-196246
15	466.5	10.6	422	12	US-10-425-114-45517

16	464.5	10.6	357	12	US-10-424-599-252964
17	459.5	10.5	346	12	US-10-425-114-57043
18	456.5	10.4	545	16	US-10-437-963-195284
19	457	10.4	427	12	US-10-424-599-193555
20	457	10.4	434	12	US-10-425-114-36844
21	452	10.3	376	16	US-10-437-963-175612
22	442	10.1	422	16	US-10-437-963-132604
23	440.5	10.0	1053	12	US-10-221-278-343
24	440.5	10.0	1053	15	US-10-291-172-343
25	439.5	10.0	483	12	US-10-424-599-284054
26	439.5	10.0	469	12	US-10-425-114-49887
27	439.5	10.0	469	12	US-10-425-114-56056
28	438.5	10.0	3993	16	US-10-437-963-168323
29	438	10.0	429	12	US-10-425-114-57924
30	438	10.0	429	12	US-10-425-114-67484
31	437.5	10.0	483	12	US-10-425-114-40870
32	436.5	9.9	919	15	US-10-108-260A-4122
33	435.5	9.9	453	12	US-10-424-599-284055
34	434.5	9.9	542	16	US-10-437-963-196087
35	433	9.9	454	12	US-10-425-114-40125
36	432	9.8	425	9	US-09-828-313-29
37	432	9.8	1330	15	US-10-108-260A-3237
38	431.5	9.8	417	12	US-10-425-114-46822
39	431	9.8	459	12	US-10-425-114-72504
40	430	9.8	604	16	US-10-437-963-173220
41	428	9.7	482	16	US-10-437-963-189054
42	427.5	9.7	367	16	US-10-250-613-6
43	426	9.6	536	16	US-10-437-963-149523
44	423.5	9.6	453	12	US-10-424-599-149508
45	423	9.6	426	12	US-10-425-114-60778

ALIGNMENTS

RESULT 1

US-09-947-199-2
; Sequence 2, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-947-199-2

Query Match 100.0%; Score 4390; DB 9; Length 835;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 835; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGNYSRPTQTCTDEWKVKVSESYVITIERLDDLOIKEKELTELNRNIFGDEAFSKVNL 60
Db 1 MGNYSRPTQTCTDEWKVKVSESYVITIERLDDLOIKEKELTELNRNIFGDEAFSKVNL 60
Qy 61 NYRTENGLSLHLCCTCGGKSHIRTLMLKGLRPSRLTNGFTALHLAVYKDNALITSL 120
Db 61 NYRTENGLSLHLCCTCGGKSHIRTLMLKGLRPSRLTNGFTALHLAVYKDNALITSL 120
Qy 121 LHSGADIQQVGYGLTALHIAITAGHLEAADVLLQGANVNIQDAVFTPLHIAAYGHE 180

Db	121	LHSGADIQVGYGGTALHTATAGHLEADVLLQGANVNIQDAVFFPLHIAAYGHE	180	Query Match	100.0%;	Score 4390;	DB 16;	Length 835;
Qy	181	QVTRLLKFGADVNVSGVGRPLHLASAKGFLNIAKLMEBSGSKADVNAQDNEDHVPJH	240	Best Local Similarity	100.0%;	Pred. No. 0;		
Db	181	QVTRLLKFGADVNVSGVGRPLHLASAKGFLNIAKLMEBSGSKADVNAQDNEDHVPJH	240	Matches	835;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	241	FCSRFGHDIIVKYLQSDLEVPQHVNIYGDTPHLACYNKGEVAKELIQSGTESLTK	300					
Db	241	FCSRFGHDIIVKYLQSDLEVPQHVNIYGDTPHLACYNKGEVAKELIQSGTESLTK	300					
Qy	301	ENIFSETAFHSACTYKGSIDLKFLDQNVININHQGRDGHGTLHSAACVGHGIRLVQFLL	360					
Db	301	ENIFSETAFHSACTYKGSIDLKFLDQNVININHQGRDGHGTLHSAACVGHGIRLVQFLL	360					
Qy	361	DNGADMNLVACDPSRSGSEKDEQTCLMWAYEKGHDAIVTLLKHYPQDELPCNEYSQPG	420					
Db	361	DNGADMNLVACDPSRSGSEKDEQTCLMWAYEKGHDAIVTLLKHYPQDELPCNEYSQPG	420					
Qy	421	GDGSYVSVPSPLGKIKSMTKEKADILLLRAGLPSHFLQLSEIEFHEIIGSGSGFKVYKG	480					
Db	421	GDGSYVSVPSPLGKIKSMTKEKADILLLRAGLPSHFLQLSEIEFHEIIGSGSGFKVYKG	480					
Qy	481	RCRNKIVAIKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLNPPSQFAIVTQ	540					
Db	481	RCRNKIVAIKRYRANTYCSKSDVDMFCREVSILCOLNHPCVIOFVGACLNPPSQFAIVTQ	540					
Qy	541	YISGGSLSLHLEQKRIIDLOSKLIIADVAKGMEYLHNLTQPIIHRDLNSHNIILYEDG	600					
Db	541	YISGGSLSLHLEQKRIIDLOSKLIIADVAKGMEYLHNLTQPIIHRDLNSHNIILYEDG	600					
Qy	601	HAVVADFGESRFLQSLDEDMNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT	660					
Db	601	HAVVADFGESRFLQSLDEDMNTKQGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILT	660					
Qy	661	GEIPFAHLKPAADAAADWYHHRPPIGYSIPKIPISILLIRGNWACPEGRPEFSEVVMKLE	720					
Db	661	GEIPFAHLKPAADAAADWYHHRPPIGYSIPKIPISILLIRGNWACPEGRPEFSEVVMKLE	720					
Qy	721	ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA	780					
Db	721	ECLCNIELMSPASSNSGSLSPSSSDCLVNRGGPGRSHVAALRSRFELEYALNARSYAA	780					
Qy	781	LSQSAGQYSSQGLSLEEMKRSLOVTPIDKYGVYSDPMSSMHFSCRNSSSFEDSS	835					
Db	781	LSQSAGQYSSQGLSLEEMKRSLOVTPIDKYGVYSDPMSSMHFSCRNSSSFEDSS	835					
RESULT 2				US-10-626-173-2				
				; Sequence 2, Application US/10626173				
				; Publication No. US20040110232A1				
				; GENERAL INFORMATION:				
				; APPLICANT: Raju, Jeyaseelan				
				; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES				
				; TITLE OF INVENTION: THEREFOR				
				; FILE REFERENCE: MNI-068CP2				
				; CURRENT APPLICATION NUMBER: US/10/626,173				
				; CURRENT FILING DATE: 2003-07-24				
				; PRIOR APPLICATION NUMBER: US/09/947,199A				
				; PRIOR APPLICATION NUMBER: US 60/111,938				
				; PRIOR FILING DATE: 1998-12-11				
				; PRIOR APPLICATION NUMBER: US 09/291,839				
				; PRIOR FILING DATE: 1998-04-14				
				; PRIOR APPLICATION NUMBER: US 09/458,457				
				; NUMBER OF SEQ ID NOS: 9				
				; SOFTWARE: PatentIn Ver. 2.0				
				; SEQ ID NO 2				
				; LENGTH: 835				
				; TYPE: PRT				
				; ORGANISM: Homo sapiens				
				US-10-626-173-2				
RESULT 3				US-09-947-199-8				
				; Sequence 8, Application US/09947199				
				; Patent No. US20020127684A1				
				; GENERAL INFORMATION:				
				; APPLICANT: Raju, Jeyaseelan				
				; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES				
				; TITLE OF INVENTION: THEREFOR				
				; FILE REFERENCE: MNI-068CP2				
				; CURRENT APPLICATION NUMBER: US/09/947,199				
				; CURRENT FILING DATE: 2001-09-05				

;; PRIOR APPLICATION NUMBER: 60/111,938
;; PRIOR FILING DATE: 1998-12-11
;; PRIOR APPLICATION NUMBER: 09/291,839
;; PRIOR FILING DATE: 1999-04-14
;; PRIOR APPLICATION NUMBER: 09/458,457
;; PRIOR FILING DATE: 1999-12-10
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 8
;; LENGTH: 835
;; TYPE: PRT
;; ORGANISM: Rattus norvegicus
;; US-09-947-199-8

Query Match 92.9%; Score 4079; DB 9; Length 835;
Best Local Similarity 91.4%; Pred. No. 2.4e-309;
Matches 763; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

Qy 1 MGNYSRPTCTDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL 60
Db 1 MGNYSRPTCTDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL 60

Qy 61 NYRTENGLSLHLCCTCGGKSHIRLMKGLRPSRLTRNGFTALHLAVYKDNALITSL 120
Db 61 NYRTENGLSLHLCCTCGGKSHIRLMKGLRPSRLTRNGFTALHLAVYKDNALITSL 120

Qy 121 LHSAGDIQVGVGGTALHIAIAGHLEAADVLLQGANVNIQDAVFTPLHIAAYYGH 180
Db 121 LHSAGDVQVGVGGTALHIAIAGHLEAADVLLQGANVNIQDAVFTPLHIAAYYGH 180

Qy 181 QVTRLLKFGADVNVSGEVDRLPLHLASAKGFNIAKLMEEGSKADVNAQDNEDHVP 240
Db 181 QVTSVLLKFGADVNVSGEVDRLPLHLASAKGFNIAKLMEEGSKADVNAQDNEDHVP 240

Qy 241 FCSRFGHDIIVKYLQSDLEVOQPHVNIYGDTPHLACVNGKFEVAKELIQTGESLTK 300
Db 241 FCSRFGHNIIVSYLLQSDLEVOQPHVNIYGDTPHLACVNGKFEVAKELIQTGESLTK 300

Qy 301 ENIFSETAFHSACTYKSIDLVKFLDDQNVININHGDRDGTGLHSACVHGHIRLVQ 360
Db 301 ENIFSETAFHSACTYKSIDLVKFLDDQNVININHGDRDGTGLHSACVHGHIRLVQ 360

Qy 361 DNGADNVLVACDPSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPQDELPCEYSPG 420
Db 361 DNGADNVLVACDPSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPQDELPCEYSPG 420

Qy 421 GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFHLQSLSEIEPHEIIGSGFGKYKG 480
Db 421 GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFHLQSLSEIEPHEIIGSGFGKYKG 480

Qy 481 RCRNKIVAKRYRANTYCSKSDVDMFCREVSVILCOLNHPCVIQFVGCINDPSPFAIV 540
Db 481 RCRNKIVAKRYRANTYCSKSDVDMFCREVSVILCOLNHPCVIQFVGCINDPSPFAIV 540

Qy 541 YISGGSLSLLHEQKRIIDLOSKLIIADVAKMEYHLNLTPIIHRDLNSHNLLEYDG 600
Db 541 YISGGSLSLLHEQKRIIDLOSKLIIADVAKMEYHLNLTPIIHRDLNSHNLLEYDG 600

Qy 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTCTRYTIKADVFSYSLCWE 660
Db 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTCTRYTIKADVFSYSLCWE 660

Qy 661 GETPFAHLKPAADAAADMAVHHIRPPIGYSPKPISSLIIRGNACPEGRPFSEVWK 720
Db 661 GETPFAHLKPAADAAADMAVHHIRPPIGYSPKPISSLIIRGNACPEGRPFSEVWK 720

Qy 721 ECLCNELMSPASSNSGSLSPSSSDCLLVNRGGPGRSHVAALRFRFEYALNARSYAA 780
Db 721 ECLCNELMSPASSNSGSLSPSSSDCLLVNRGGPGRSHVAALRFRFEYALNARSYAA 780

Qy 781 LSQSAQYSSQGLSLEMKRSIQYPIDKYGVSPDPMGSMHPSCRNSSFEDSS 835
Db 781 WSQSVGTHNPGLSLEMKRSIQYSTVDKYGVYSPDPMGLTHLSRQDSDNFEDSN 835

RESULT 4
US-10-626-173-8
;; Sequence 8, Application US/10626173
;; Publication No. US20040110232A1
;; GENERAL INFORMATION:
;; APPLICANT: Raju, Jeyaseelan
;; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
;; TITLE OF INVENTION: THEREFOR
;; FILE REFERENCE: MNI-068CP2
;; CURRENT APPLICATION NUMBER: US/10/626,173
;; CURRENT FILING DATE: 2003-07-24
;; PRIOR APPLICATION NUMBER: US/09/947,199A
;; PRIOR FILING DATE: 1998-12-11
;; PRIOR APPLICATION NUMBER: US/09/291,839
;; PRIOR FILING DATE: 1999-04-14
;; PRIOR APPLICATION NUMBER: US/09/458,457
;; PRIOR FILING DATE: 1999-12-10
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 8
;; LENGTH: 835
;; TYPE: PRT
;; ORGANISM: Rattus norvegicus
;; US-10-626-173-8

Query Match 92.9%; Score 4079; DB 16; Length 835;
Best Local Similarity 91.4%; Pred. No. 2.4e-309;
Matches 763; Conservative 38; Mismatches 34; Indels 0; Gaps 0;

Qy 1 MGNYSRPTCTDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL 60
Db 1 MGNYSRPTCTDEWKKVSESYVITIERLEDDLOIKEKELTELNRNIFGSDFAFSKYNL 60

Qy 61 NYRTENGLSLHLCCTCGGKSHIRLMKGLRPSRLTRNGFTALHLAVYKDNALITSL 120
Db 61 NYRTENGLSLHLCCTCGGKSHIRLMKGLRPSRLTRNGFTALHLAVYKDNALITSL 120

Qy 121 LHSAGDIQVGVGGTALHIAIAGHLEAADVLLQGANVNIQDAVFTPLHIAAYYGH 180
Db 121 LHSAGDVQVGVGGTALHIAIAGHLEAADVLLQGANVNIQDAVFTPLHIAAYYGH 180

Qy 181 QVTRLLKFGADVNVSGEVDRLPLHLASAKGFNIAKLMEEGSKADVNAQDNEDHVP 240
Db 181 QVTSVLLKFGADVNVSGEVDRLPLHLASAKGFNIAKLMEEGSKADVNAQDNEDHVP 240

Qy 241 FCSRFGHDIIVKYLQSDLEVOQPHVNIYGDTPHLACVNGKFEVAKELIQTGESLTK 300
Db 241 FCSRFGHNIIVSYLLQSDLEVOQPHVNIYGDTPHLACVNGKFEVAKELIQTGESLTK 300

Qy 301 ENIFSETAFHSACTYKSIDLVKFLDDQNVININHGDRDGTGLHSACVHGHIRLVQ 360
Db 301 ENIFSETAFHSACTYKSIDLVKFLDDQNVININHGDRDGTGLHSACVHGHIRLVQ 360

Qy 361 DNGADNVLVACDPSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPQDELPCEYSPG 420
Db 361 DNGADNVLVACDPSRSGEKDEQTCIMWAYEKGHDAIVTLLKHYPQDELPCEYSPG 420

Qy 421 GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFHLQSLSEIEPHEIIGSGFGKYKG 480
Db 421 GDGSYVSPPLGKIKSMTEKADILLRAGLPSHFHLQSLSEIEPHEIIGSGFGKYKG 480

Qy 481 RCRNKIVAKRYRANTYCSKSDVDMFCREVSVILCOLNHPCVIQFVGCINDPSPFAIV 540
Db 481 RCRNKIVAKRYRANTYCSKSDVDMFCREVSVILCOLNHPCVIQFVGCINDPSPFAIV 540

Qy 541 YISGGSLSLLHEQKRIIDLOSKLIIADVAKMEYHLNLTPIIHRDLNSHNLLEYDG 600
Db 541 YISGGSLSLLHEQKRIIDLOSKLIIADVAKMEYHLNLTPIIHRDLNSHNLLEYDG 600

Qy 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTCTRYTIKADVFSYSLCWE 660
Db 601 HAVVADFGESRFLQSLDEDNMTKQGNLRWMAPEVFTCTRYTIKADVFSYSLCWE 660

Db 601 HAVVADGESRFQSLDEDNMTKOPGNLRWMAEVFTQCTRTYTIKADVFYSILCWELLT 660
Qy 661 GEIPFAHLKPAADAAADWAYHRRPPIGYSIPKPSISLLIRGNWACPEGRPESEVVMKLE 720
Db 661 GEIPFAHLKPAADAAADWAYHRRPPIGYSIPKPSISLLIRGNWACPEGRPESEVVMKLE 720
Qy 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGPGGRSHVAALRSRFELEYALNARSAA 780
Db 721 ECLCNIELMSPASSNSGSLSPSSSDCLVNRGPGGRSHVAALRSRFELEYALNARSAA 780
Qy 781 LSQAGQYSSOGLSLEEMKRSLOVTPIDKYGYVSDPMSSMHFSCNRSNSSFDDSS 835
Db 781 WSQSVGTHSNPGLSLEEMRSTQYSTVDKYGYVSDPMSLTHLSRQDDSNFSDSN 835

RESULT 5

US-10-664-421-138
; Sequence 138, Application US/10664421
; Publication No. US20040142864A1

GENERAL INFORMATION:
; APPLICANT: BREMER, RYAN
; APPLICANT: IERAHIM, PRAEHA
; APPLICANT: KUMAR, ABHINAV
; APPLICANT: MANDIYAN, VALSAN
; APPLICANT: MILBURN, MICHAEL V.
; TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE
; FILE REFERENCE: 039363/0703
; CURRENT APPLICATION NUMBER: US/10/664,421

; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/412,341
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/411,398
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-664-421-138

Query Match 43.1%; Score 1893; DB 16; Length 359;
Best Local Similarity 100.0%; Pred. No. 3.4e-139; Mismatches 0; Indels 0; Gaps 0;
Matches 359; Conservative 0;

Qy 413 CNEYSQGGDGSYVSPPLGKIKSMTEKADILLRAGLPSHFHLQLSIEFHEIIGSG 472
Db 1 CNEYSQGGDGSYVSPPLGKIKSMTEKADILLRAGLPSHFHLQLSIEFHEIIGSG 60
Qy 473 SFGKYKGRCKNKIVAKRYRANTYCSKSDVDMFCREVSIILCOLNHPCVIQVGCINDP 532
Db 61 SFGKYKGRCKNKIVAKRYRANTYCSKSDVDMFCREVSIILCOLNHPCVIQVGCINDP 120
Qy 533 SOFAIVTQYISGGSLFSLHHEQKRILDIQSKLIIVADVAKGMEYHLNLTQPIIHRDLNSH 592
Db 121 SOFAIVTQYISGGSLFSLHHEQKRILDIQSKLIIVADVAKGMEYHLNLTQPIIHRDLNSH 180
Qy 593 NTLLEDGHAVVADGESRFLQSLDEDNMTKOPGNLRWMAEVFTQCTRTYTIKADVFSA 652
Db 181 NTLLEDGHAVVADGESRFLQSLDEDNMTKOPGNLRWMAEVFTQCTRTYTIKADVFSA 240
Qy 653 LCLWEILLTGEPFAHLKPAADAAADWAYHRRPPIGYSIPKPSISLLIRGNWACPEGRPEF 712
Db 241 LCLWEILLTGEPFAHLKPAADAAADWAYHRRPPIGYSIPKPSISLLIRGNWACPEGRPEF 300
Qy 713 SEVVMKLECLCNIELMSPASSNSGSLSPSSSDCLVNRGPGGRSHVAALRSRFELEY 771
Db 301 SEVVMKLECLCNIELMSPASSNSGSLSPSSSDCLVNRGPGGRSHVAALRSRFELEY 359

RESULT 6

US-10-437-963-195691

; Sequence 195691, Application US/10437963
; Publication No. US20040123343A1

GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195691
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Oryza sativa

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91615C.1.pap
US-10-437-963-195691

Query Match 11.1%; Score 486.5; DB 16; Length 359;
Best Local Similarity 37.0%; Pred. No. 2.1e-29; Mismatches 105; Indels 37; Gaps 9;
Matches 117; Conservative 57;

Qy 468 IIGSGFGKYKGRCKNKIVAKRYRANTYCSKSDVDMFCREVSIILCOLNHPCVIQVGA 527
Db 49 MIASGSGDLYHGYLGEDVAVKILASE-HLNKNVWNEFTQEVILREVQHTNVVRFIGA 107
Qy 528 CLNDPSQFAIVTQYISGGSLFSLHHEQKRILDIQSKLIIVADVAKGMEYHLNLTQPIIHR 587
Db 108 C-TKPPQFCITTEYMSGGSLYDFVHRQHNVLDTLLKFAVDVCRGMCYLLHQ--RGIHR 164
Qy 588 DLNHNILLYEDGHAVVADGESRFLQSLDEDNMTKOPGNLRWMAPE-VFTQCT----- 640
Db 165 DLKSNLLMDKDHVAVDFGVARFQD--QGGNMTAETGTYRWMAPEHFFSPCSLGERMN 222
Qy 641 -----RYTIKADVFYSALCLWEILLTGEPFAHLKPAADAAADWAYHRRPPIGYS 689
Db 223 WDKKEIVINHQPYDNKADVFSPAIVLWELITSKIPTYNTMTPLQAAVGVV--QGLRFGLPEN 281
Qy 690 IPRPISLLIRGNWACPEGRPESEVVMKLEELCNIELMSPASS-----NSSGSLSPSS 744
Db 282 AHPOLLDMRCWEGIPSNRPPFDILAEILLARVQEQEHSNQMKHNSNLK--- 338
Qy 745 SSDCLVNRGPGGRSHV 760
Db 339 ---LSNAGMLGSNYI 350

RESULT 7

US-10-424-599-271998
; Sequence 271998, Application US/10424599
; Publication No. US20040031072A1

GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 271998
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Glycine max

```
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(376)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87633C.1.p.p
US-10-424-599-271998

Query Match      11.0%; Score 483.5; DB 12; Length 376;
Best Local Similarity 38.8%; Pred. No. 3.9e-29;
Matches 112; Conservative 54; Mismatches 108; Indels 15; Gaps 7;

Qy 456 FHLQSLBIEHEIIGSGFGKVKGRCKNKIYAIKRYRANTYCSKSDVDM---FCREYSI 512
Db 93 WEIDTQWKYENKVGSGFDLYRGTCYQDVAIKVLP-----ERISTDLREFAEQYVI 148
Qy 513 LCOLNHPCVIQFGACLNDFSPFAIVTOYISGSLFSLHKEQKRLDLOSKLIIADVAK 572
Db 149 MRKIRHNKVVOFQIGACTRPN-LCIVTEFMSRGLYDFLHKQGVFKLPSSLKVAIDVSK 207
Qy 573 GMEYLNLTQPIIHRDLNHNILLYEDGHAVVADFGESRFLQSLDEDNMTKQGNLRMWA 632
Db 208 GMYLHQ--NNIHRDLKTANLLMDENGVKADFGVARV--QTSQGVMTAETGYRMA 263
Qy 633 PEVFTQCTRYTIKADVFSALCLWEILTGEIPFAHLKPAADAAADMAVHHRIPPIGYSIPK 692
Db 264 PEVI-EHKPDQKADVFSFGIALWELLTGELPYSCITPLQAAVGVVQKGLRPTIPKNTHP 322
Qy 693 PISGLLRGNACPEGRPEFSEVVMKLECLNCLNIELMSPASSNSGSL 741
Db 323 RLSELLQRCWQDPTORPNEFSEIILQOTA--KEYNDHKDKSSHGFLS 369

RESULT 8
US-10-437-963-156923
; Sequence 156923, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 156923
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56544C.1.p.p
US-10-437-963-156923

Query Match      11.0%; Score 482; DB 16; Length 603;
Best Local Similarity 29.4%; Pred. No. 9.8e-29;
Matches 130; Conservative 83; Mismatches 189; Indels 40; Gaps 10;

Qy 296 ESLTKENIFSETAFHSACTYKSIDLVKFLLDONVININHQGRD-----GHTGLHSA 347
Db 154 EEVASRLNLRQSHPPFPAFGSSTNLEALALEAS---KSQGDHSDSDNVNRPMEI 209
Qy 348 CYHGHIR-----LVQFLLDNGADMLVACDPSSSGEKDEQTCMLMAYEKGHDAIVILL 401
Db 210 TFSIDTRPKLISELTSULGELGNICBAHAFSTNDGYSLDFVVGWGHDETFDLIESVR 269
Qy 402 KHYKRPQDELPCNEYSQPGDGSVSVPSPLGKIKMTKEKADILLRAGLPSHFHLQLS 461

; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(376)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87633C.1.p.p
US-10-424-599-271998

Query Match      11.0%; Score 483.5; DB 12; Length 376;
Best Local Similarity 38.8%; Pred. No. 3.9e-29;
Matches 112; Conservative 54; Mismatches 108; Indels 15; Gaps 7;

Qy 456 FHLQSLBIEHEIIGSGFGKVKGRCKNKIYAIKRYRANTYCSKSDVDM---FCREYSI 512
Db 93 WEIDTQWKYENKVGSGFDLYRGTCYQDVAIKVLP-----ERISTDLREFAEQYVI 148
Qy 513 LCOLNHPCVIQFGACLNDFSPFAIVTOYISGSLFSLHKEQKRLDLOSKLIIADVAK 572
Db 149 MRKIRHNKVVOFQIGACTRPN-LCIVTEFMSRGLYDFLHKQGVFKLPSSLKVAIDVSK 207
Qy 573 GMEYLNLTQPIIHRDLNHNILLYEDGHAVVADFGESRFLQSLDEDNMTKQGNLRMWA 632
Db 208 GMYLHQ--NNIHRDLKTANLLMDENGVKADFGVARV--QTSQGVMTAETGYRMA 263
Qy 633 PEVFTQCTRYTIKADVFSALCLWEILTGEIPFAHLKPAADAAADMAVHHRIPPIGYSIPK 692
Db 264 PEVI-EHKPDQKADVFSFGIALWELLTGELPYSCITPLQAAVGVVQKGLRPTIPKNTHP 322
Qy 693 PISGLLRGNACPEGRPEFSEVVMKLECLNCLNIELMSPASSNSGSL 741
Db 323 RLSELLQRCWQDPTORPNEFSEIILQOTA--KEYNDHKDKSSHGFLS 369

RESULT 9
US-10-425-114-40404
; Sequence 40404, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40404
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-022-C7_FLI.p.p
US-10-425-114-40404

Query Match      10.9%; Score 480.5; DB 12; Length 478;
Best Local Similarity 27.4%; Pred. No. 9.3e-29;
Matches 170; Conservative 95; Mismatches 179; Indels 177; Gaps 29;

Qy 145 GHLEAADVLLQHGANNVIOQAVFFFTPLHIAAYGHEQVTRLLKFGADVNVSGEVDRLP 204
Db 7 GYAEASD-----LEFRD-----RLAAHFG-----RLPHSYQLDINVD-KASDLV 45
Qy 205 H---LASAKG-----FLNIAKLMEEGSKADYNAQDNEDHVPFLHFCRSRFGH-HDI 250
Db 46 HONVLAELAKDPDRRPAFYVRFRLIEDMSAYSDASEEGDDGDDLSVRQDTPYTHIHEI 105
Qy 251 VKYLLQSDLEVOQHVNVNYIGDTPHLACYNKGFVEKELIIOISGTES-----LTKENIFS 305
Db 106 VFSTI-----DKP-----KLSQSLALLSDIGLNIREAHVFS 137
Qy 306 ETAFHSACTYKSIDLVKFLLDONVININHQGRDGTGLHSAFYEGHIRLVQFLLDNGAD 365
Db 138 T---HD---GYSLDV--FVVDGWPV---QDAD---GLHKA-----163
Qy 366 MNLVACDPSSSGEKDEQTCMLMAYEKGHDAIVTLKHYKRPQDELPCNEYSQPGDGSY 425
Db 164 ---LEASILRNEGS-----WSSG-----SHSSAAERTLP---FOVKGGEM-- 197
```



```

; SEQ ID NO 67647
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3354-074-A5_FLI.pep
US-10-425-114-67647

Query Match 10.8%; Score 473; DB 12; Length 462;
Best Local Similarity 38.7%; Pred. No. 3.4e-28;
Matches 108; Conservative 57; Mismatches 96; Indels 18; Gaps 8;

QY 468 ITIGSGFGKVKGRGNKVAIKRYRANTYCSKSDVDMFCREYSILCOLNHPCVIQFVGA 527
DB 193 WVASGSGDLXGYTGEDVAVKVIKRAE-HLNKNVWNETQEVYILREVOHKNVRFIGA 251
QY 528 CLNDPSQFAIVTYQISGSLPSLHHEKRLIDQSKLIITADVAKMEYLNHLTOPIIHR 587
DB 252 C-TKPPQPCFIITEYMSGSLYDFVHKQHNLNLTLLKFAVDVCRGMCYLHE--RGIHR 308
QY 588 DLNLSNILLYEDGHAVADFCESRFQLQSLDEDN--MTQPGNLRMAPEVFTQCTRTYI 644
DB 309 DLKTANLMDKHVVKVADFGVARP-----QDQGVMTAETGYRMAPEVINH-QPYDN 362
QY 645 KADVFYSALCLWEILTGBEIPFAHLKPAAAAADWAYHHIRPPIGYSPKPISSLIRGWA 704
DB 363 KADVFESFAIVWELITSKIPIYESMTPLQAAVGR-QGLRPLPKTHPKVLDLMQRCWEA 421
QY 705 CPBGEPPESEVVMKLECLCNIE----LMSFASNSGS 739
DB 422 DFSARPAFPDILAELEDLLAQVQGTGKTAHAPSNSNT 460

RESULT 13
US-10-425-114-64160
; Sequence 64160, Application US/10425114
; Publication NO. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Jiu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64160
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3592-081-E1_FLI.pep
US-10-425-114-64160

Query Match 10.8%; Score 472.5; DB 12; Length 338;
Best Local Similarity 37.0%; Pred. No. 2.4e-28;
Matches 108; Conservative 58; Mismatches 113; Indels 13; Gaps 6;

QY 432 LKIKSMYKEXADILLRAGIPSHFHLQLSEIPEHFIIGSGSGKVKYKGCERNKIVAIR 491
DB 26 LKXMQIAENSAADRVEIPTDGASEWEDIVLLKFGNKVAGSYGDLYRGTYCSQDVAIKV 85
QY 492 YRANTYCSKSDVDM--FCREYSILCOLNHPCVIQFVGACLNDPDSQFAIVTYQISGSLF 548
DB 86 LXP-----ERINADMQREFAQEVYIMRKVRHKVQVFIGACTKFPN-LCIVTFEMSGSVY 140
QY 549 SILHSGKRLIDQSKLIITADVAKMEYLNHLTOPIIHRDLNLSNILLYEDGHAVADFG 608
DB 141 DYLNHKHKGVKFLPALVGVATDVSKGMSYLHO--NNIIHRDLKTANLMDQENGTVKVADFG 198

```

```

Qy 609 ESFLOSLEDNNTKOPGNLRWMADEVFTQCTRYTIKADVSYALCLWEILLTGEIPIPAHL 668
Db 199 VAEV--KAOSGVNMTAETGYRWMAPEVI-EHRPYDHKADVFSGILMWELLTKGPIPEYL 255
Qy 669 KPAAAADNAVYHHIRPIGVSIPKPISSILLIRGNWACPEGRPEFSEVVMKLE 720
Db 256 TPLQAVGVVQKGLRTIPKHTHAKLSSELLQKWCQODPTQRPDESEILETLQ 307

RESULT 14
US-10-424-599-196246
; Sequence 196246, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 196246
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(551)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_19237C.1.pep
US-10-424-599-196246

```

```

Query Match      10.7%; Score 468; DB 12; Length 551;
Best Local Similarity 33.9%; Pred. No. 1.1e-27;
Matches 122; Conservative 64; Mismatches 148; Indels 26; Gaps 8;

QY      381 DBOTCLMWAYEKGHDAIVTLTKHKVRQDDELPNCNEYSQPGDGSVSVSPSLGKTKSMTK 440
Db      209 ETEKELKAALRG-----VLKIERQVKSSP-----QSVSSVDKP---DQAGWK 249

QY      441 EKADILLRLAGLSPHFHLOQSEIEFHEIIGSGFGKYVGRCRNKIVAKRYRANTYCSK 500
Db      250 SELDYLTPTDGTDWEDPKHLKYGTQASGSYGELFGVYCSOEVAIKVLKADHVNSE 309

QY      501 SDVDMFECEVSILCOLNHPCVIOFVGACLANDPSQFAIVTCYISGGSLFSLHEOKRILDL 560
Db      310 LQRE-FAQEVIMRKVRHKNVQPIGACTKSP-RLCIVTFEFGSGSVGYLHKQGFKEF 367

QY      561 QSKLIIVADVAKGMEYLNHLNLTQPIIHRDLNSHNLILEYDGHAVADFGESRFLQSLDEBN 620
Db      368 PTLKVAIDVSGKNYLLHQ--HNIHEDLKAALLNDENCTKVADFGVARV--KAQSGV 423

QY      621 MTKQGNRLRWMAPEVFTQCTRYTIKADVFYALCLWEILITGEIPAHLKPAAAAADYAH 680
Db      424 MTAETGTGRWMAPEVI-EHKFYDHKADVFSGVLVWELLTGLPYEYITPLQAATGVVQK 482

QY      681 HIRPPIGVSIKPISSILLIRGNWACPEGRSEFSEVVMKLEBCLCHIELMSPASSNSSGL 740
Db      483 GURPTIPKNTHPKFVELLERSWOODPTLRDPFSEIETILQOLAKEVGDEERHKDKSGL 542

RESULT 15
US-10-425-114-45517
; Sequence 45517, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua

```

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 45517
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700243456_FLI.pep
US-10-425-114-45517

Query Match          10.6%; Score 466.5; DB 12; Length 422;
Best Local Similarity 32.9%; Pred. No. 9.7e-28;
Matches 131; Conservative 62; Mismatches 142; Indels 63; Gaps 15;

QY 354 RLQVQLLDNGADMLIVAC-----DPSRSSGGEKDEQTCIMWAYEKGHDAIVTLKHYKRPQ 408
Db 9 RRARELL--SISTNPIACHKYCTDASHCFNCRSIQTL-----HP---RKLVRSGDM 55

QY 409 DELPCNEYSQGGDGSYVSPSP-----LGKIKSMTEKADILL-----LRAGLPS 454
Db 56 KRLQCFKSSGGGNGG-----GPGKRLERLSLGEYKKAWSKYLIVAPPGARIRSGGEE 110

QY 455 HFHLQLSEIEFHEIIGSGFGKVKYGRCKNKIVAIKRYRANTYCSKSDVD-----M 505
Db 111 LWSADLSKLEIRGFASGRHSRVYSGRYTGREVAIK-----WVSGPEEDALAAELERQ 164

QY 506 FCREVSIQLQNLHPCVIOFVGACLNDSQFAIVTQYISGGSLSLLEQK-RILDLSKL 564
Db 165 FASEVALLRLHHPNIIISFVAAC-KKPPVFCIITEFMAGGSLRKYLHQEPHSVPLNLVL 223

QY 565 IIAVDVAKGMEVYLNLTQPIIHRDLNSHNILLVEDGHAVVADFGESRFLQSLDEDNMTKQ 624
Db 224 KLALDIARKMSYLH--SGILHRLKSENILLGEDMSVKVADFG-----ISCLSQCGSK 277

QY 625 --PGNLRWMAPEVFTQCTRYTIKADVFSYALCLWEILTGPIFAHLKPAAADYAHYHI 682
Db 278 GFTGYRWAPFEMIKE-EHHTRKVDVYSFGIVMWEILTALVPFSDMTPEQAAVAVALKVA 336

-QY 683 RPPIGYSIPKPISSILLINGWACPEGRPEFSEVVMKLE 720
Db 337 RPPLPASCFAISHLMQCNATNPDKRPQFDDIVAILE 374
```

Search completed: September 2, 2004, 14:36:20
Job time : 146 secs